

UNIVERSIDADE ESTADUAL DE SANTA CRUZ
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR



**TcSERPIN, um inibidor proteico que interage com cisteíno-protease de folhas
do cacaueiro**

MONALIZA MACÊDO FERREIRA

ILHÉUS – BAHIA – BRASIL

Setembro de 2018

MONALIZA MACÊDO FERREIRA

**TcSERPIN, um inibidor proteico que interage com cisteíno-protease de folhas
do cacaueiro**

Dissertação apresentada à Universidade Estadual de Santa Cruz como parte das exigências para obtenção do título de Mestre em Genética e Biologia Molecular.

Área de concentração: Genética e Biologia Molecular

ILHÉUS – BAHIA – BRASIL

Setembro de 2018

MONALIZA MACÊDO FERREIRA

**TcSERPIN, um inibidor proteico que interage com cisteíno-protease de folhas
do cacauero**

Dissertação apresentada à Universidade Estadual de Santa Cruz como parte das exigências para obtenção do título de Mestre em Genética e Biologia Molecular.

Área de concentração: Genética e Biologia Molecular

APROVADA:

Dr^a Jannaina Velasques da Costa Pinto
UFSB

Dr. Márcio Gilberto Cardoso Costa
UESC

Dr^a Milena do Amaral Santos
UESC

Dr. Carlos Priminho Pirovani
UESC – orientador

AGRADECIMENTOS

Agradeço acima de todas as coisas a Deus, por sempre estar ao meu lado, em cada passo e a cada momento de minha breve existência.

À minha mãe, por todo incentivo e dedicação em todos esses anos.

Ao Programa de Pós Graduação em Genética e Biologia Molecular da Universidade Estadual de Santa Cruz e a CAPES, pela concessão da bolsa de pesquisa.

A todos os meus companheiros e amigos que estiveram presente durante a minha jornada: à Geiseane Veloso Amaral e Maria Luíza, minhas parceiras de laboratório; à Maria Zugaib, por toda a parceria e orientação desde a época de iniciação científica; à Moara, por toda a gentileza e humildade em seus ensinamentos; à Keilane, Juliano Santana e Dona Jô, pelo carinho; e a todos os meus colegas companheiros da turma de 2017, muito obrigada, pelos momentos de alegria.

Meus agradecimentos especiais à minha grande parceira e amiga, Ákylla Maria Martins Alves, por toda a instrução, paciência, incentivo, por toda a ajuda, e, acima de tudo, pela amizade. E ao meu orientador, Dr. Carlos Priminho Pirovani, por compartilhar o conhecimento e por toda a confiança depositada nesses anos de parceria. Sinto muito orgulho de fazer parte do grupo de pesquisa “Proteômicos” e de toda a família CBG.

EXTRATO

MACÊDO, Monaliza Ferreira. Universidade Estadual de Santa Cruz, Ilhéus, setembro de 2018. **TcSERPIN, um inibidor proteico que interage com cisteíno-protease de folhas do cacau.** Orientador: Prof. Dr. Carlos Priminho Pirovani. Co-orientadores: Profa. Dra. Virgínia Lúcia Fontes Soares e Prof. Dr. Bruno Silva Andrade.

Palavras-chave: *Theobroma cacao*, inibidores, proteases, expressão heteróloga.

Serpinas são uma superfamília de proteínas com capacidade inibitória para serino- e cisteíno-peptidases. Em humanos, o principal alvo são as serino-proteases da família da tripsina ou quimotripsina. Em plantas seu potencial inibitório, aparentemente, tem sido melhor descrito contra cisteíno-peptidases. As serpinas inibitórias possuem um padrão estrutural similar em sua principal região de sítio ativo, o “*Reactive Center Loop*” (RCL), por onde ocorre as interações com as peptidases alvo. Elas estão envolvidas em diversos processos biológicos de regulação celular, como coagulação, controle da pressão arterial, supressão gênica, regulação fisiológica e apoptose. Em plantas, estudos com trigo, cevada, arroz, abóbora, tomate, pepino e *Arabidopsis thaliana* têm sido relatados. No entanto, apesar do evidente potencial inibitório desses inibidores, nenhuma via metabólica em vegetais tem sido bem caracterizada. Sendo assim, o presente estudo objetivou caracterizar uma serpina de *Theobroma cacao*. As análises foram realizadas *in silico* e *in vitro*. A proteína denominada TcSERPIN teve sua sequência primária obtida a partir do banco de dados do genoma do cacau (cocoaGen DB). Foram realizados alinhamentos da proteína com outras sequências homólogas de serpinas de plantas e humanos com o Clustal Omega, e a partir desses alinhamentos foi gerado um dendrograma segundo o Mega 7. A estrutura 3D da proteína foi predita por meio de modelagem por homologia no Swiss Model, usando como modelo o *template* 3LE2 de *Arabidopsis thaliana*. Análises de *docking* foram realizadas em três etapas: com uma heparina de sódio obtida no PubChem; com uma tripsina de *Bos taurus* (1S0Q PDB) e uma papaína de *Carica papaya* (1QVZ PDB), ambas obtidas a partir do banco de dados PDB; e duas papaínas do cacau, modeladas por homologia a partir do modelo 5EF4 PDB de *Ambrosia artemisiifolia* e 3QJ3 PDB de *Tenebrio molitor*, a RD21 e a RD19, respectivamente. As análises de *docking* foram realizadas com os programas ClusPro e AutoDock vina e a visualização das estruturas foram obtidas com os programas Discovery Studio 4.0 e PyMol. As análises *in vitro* foram obtidas a partir da expressão da ORF sintética inserida no vetor pET-28a e a indução em células de *E. coli*, Rosetta (DE₃). Após a obtenção da proteína recombinante purificada foram realizadas análises para determinação da capacidade inibitória da serpina contra a tripsina do pâncreas de *Sus scrofa* e a papaína do látex de *Carica papaya*, utilizando o substrato BApNA, e captura de proteases de folha do cacau com a TcSERPIN imobilizada em resina CNBr-Sepharose. A estrutura secundária e a estabilidade térmica foram analisadas por espectroscopia de dicroísmo circular (CD). Os resultados mostram que TcSERPIN possui 390 resíduos de aminoácidos. Os alinhamentos demonstram que a proteína possui alta conservação dos resíduos de aminoácidos, e o dendrograma mostrou que a serpina de plantas possui motivos altamente conservados que as separam

de outros organismos. O perfil do gel SDS-PAGE mostrou a presença de uma proteína de 42,5 kDa purificada a partir da fração insolúvel do extrato bacteriano. As análises do espectro da proteína por dicroísmo circular determinou a presença de folhas- α , folhas- β e alças. As curvas de *unfolding* da proteína mostraram temperatura de *melting* de 56 a 78 °C em pH 5,4 e 7,0 respectivamente. Os testes de atividade e as análises de *docking* mostraram que TcSERPIN possui alta afinidade com a heparina e com proteases do tipo quimotripsina, tripsina e papaínas. Os resultados sobre o potencial de captura das proteases pela rTcSERPIN imobilizada em resina CNBr-Sepharose foram avaliados a partir de gel de zimograma contendo gelatina, onde foi observado que rTcSERPIN interage com peptidases do tipo cisteína da via endógena de *T. cacao*. Os resultados obtidos até o momento asseguram que TcSERPIN é uma proteína inibitória e com maior afinidade para proteases do tipo papaínas.

ABSTRACT

MACÊDO, Monaliza Ferreira. Universidade Estadual de Santa Cruz, Ilhéus, March 2016. **TcSERPIN, a proteic inhibitor that Interacts with cysteine-protease from leaves of the cocoa tree.** Advisor: Prof. Dr. Carlos Priminho Pirovani. Advisor Committee Members: Profa. Dra. Virgínia Lúcia Fontes Soares and Prof. Dr. Bruno Silva Andrade.

Key words: TcSERPIN, inhibitors, proteases, heterologous expression

Serpins make up a superfamily of proteins with inhibitory capacity for serine- and cysteine-peptidases. In humans, the preferred inhibition target is the serine proteases of the trypsin or chymotrypsin family. In plants their inhibitory potential has apparently been better described against cysteine-peptidases. Inhibitory serpins have a similar structural pattern in their main active site region, the Reactive Center Loop (RCL), where interactions with target peptidases occur. They are involved in several biological processes of cell regulation, such as coagulation, blood pressure control, gene suppression, physiological regulation and apoptosis. In plants, studies with wheat, barley, rice, pumpkin, tomato, cucumber and *Arabidopsis thaliana* have been reported. However, despite the evident inhibitory potential of these inhibitors, no metabolic pathway in plants has been well characterized. Thus, the present study aimed to characterize a serpin of *Theobroma cacao*. The analysis was performed *in silico* and *in vitro*. The protein termed TcSERPIN had its primary sequence obtained from the cacao genome database (cocoaGen DB). Alignments of the protein with other homologous sequences of plant and human serpins with Clustal Omega were performed, and from these alignments a dendrogram according to Mega 7 was generated. The 3D structure of the protein was predicted by homology modeling in the Swiss Model, using as template the 3LE2 template of *Arabidopsis thaliana*. *Docking* analyzes were performed in three steps: with a sodium heparin obtained from PubChem; with a *Bos taurus* trypsin (1S0Q PDB) and a *Carica papaya* papain (1QVZ PDB), both obtained from the PDB database; and two cacao papayas, modeled by homology from the model 5EF4 PDB of *Ambrosia artemisiifolia* and 3QJ3 PDB of *Tenebrio molitor*, RD21 and RD19, respectively. The *docking* analyzes were performed with the ClusPro and AutoDock vina programs and the visualization of the structures was obtained with the programs Discovery Studio 4.0 and PyMol. *In vitro* analyzes were obtained from expression of a synthetic ORF inserted into the pET-28a vector and induction into *E. coli*, Rosetta (DE3) cells. After the purified recombinant protein was obtained, analyzes were performed to determine the inhibitory capacity of the serpin against the trypsin of *Sus scrofa* pancreas and the papain of the *Carica papaya* latex, using the BApNA substrate, and capture of cocoa leaf proteases with the TcSERPIN immobilized on CNBr-Sepharose resin. Secondary structure and thermal stability were analyzed by circular dichroism (CD) spectroscopy. The results show that TcSERPIN has 390 amino acid residues. The

alignments show that the protein has high conservation of amino acid residues, and the dendrogram showed that plant serpins have highly conserved motifs that separate them from other organisms. The SDS-PAGE gel profile showed the presence of a purified 42.5 kDa protein from the insoluble fraction of the bacterial extract. The analysis of the protein spectrum by circular dichroism presented the presence of α -sheets, β -sheets and loops. The *Unfolding* curves of the protein showed melting temperature of 56 to 78 °C at pH 5.4 and 7.0 respectively. Activity tests and docking analyzes showed that TcSERPIN has high affinity for heparin and chymotrypsin, trypsin and papain proteases. Results on the trapping potential of proteases by rTcSERPIN immobilized on CNBr-*Sepharose* resin were evaluated from gelatin-containing zymography gel, where it was observed that TcSERPIN interacts with cysteine-like peptidases of the endogenous pathway of *T. cocoa*. The results obtained so far ensure that TcSERPIN is an inhibitory protein and with greater affinity for papain type proteases.

ÍNDICE

EXTRATO	v
ABSTRACT	vi
LISTA DE FIGURAS	xi
1 INTRODUÇÃO	12
2 OBJETIVOS	16
2.1 Objetivo geral.....	16
2.2 Objetivos específicos.....	16
3. REFERENCIAL TEÓRICO	17
3.1 Proteases.....	17
3.2 Mecanismo catalítico das preteases	19
3.3 Inibidores proteico de proteases	20
3.3.1Serpinas.....	22
3.3.1.1Mecanismos estruturais e funcionais das serpinas.....	23
3.3.2 Serpinas e afinidade com pentassacarídeos.....	25
3.3.3 Serpinas de plantas	27
3.4 Serpinas em vias de morte celular programada	29
4 MATERIAL E MÉTODOS	31
4.1 Análise <i>in silico</i> da sequência primária de TcSERPIN.....	31
4.2 Modelagem por homologia de TcSERPIN e cisteino-proteases do cacau.....	32
4.3 Expressão e purificação de TcSERPIN recombinante.....	33
4.4 Análise de Dicroísmo Circular (CD).....	34
4.5 Ensaio da atividade inibitória de TcSERPIN contra tripsina de <i>Sus scrofa</i> e papaína de <i>Carica papaya</i>	35
4.6 Análise de biologia de sistemas.....	35
4.7 Docking molecular.....	36
4.7.1 <i>Docking</i> entre TcSERPIN e as proteases do tipo serino e cisteíno proteases.....	36
4.7.2 Docking entre TcSERPIN e a heparina de sódio.....	37
4.8 Extração de proteínas nativas totais das folhas de <i>T. cacao</i>	37
4.9 Armadilha da protease por CNBr-Sepharose.....	38
5. RESULTADOS	39
5.1 Análise <i>in silico</i> da sequência primária de TcSERPIN.....	39

5.2 Alinhamento de TcSERPIN e sequências homólogas.....	39
5.3 Modelagem.....	42
5.4 Expressão e purificação de TcSERPIN recombinante.....	45
5.5 Análise da estrutura secundária de TcSERPIN em Dicroísmo Circular (CD).....	46
5.6 Análise da atividade inibitória de rTcSERPIN contra tripsina suína e a papaína de <i>Carica papaya in vitro</i>	47
5.7 Análise <i>in silico</i> da interação de TcSERPIN, com proteases do tipo serino e cisteíno	48
5.8 Rede de interação da TcSERPIN.....	50
5.9 Análise <i>in silico</i> da interação de TcSERPIN com cisteíno-proteases de <i>Theobroma cacao</i>	51
5.10 Análise da interação da heparina de sódio com TcSERPIN em sítios similares da Antitrombina-III.....	53
5.11 Busca de alvo da TcSERPIN em extrato de folhas do cacau.....	54
6. DISCUSSÃO.....	56
7. CONCLUSÃO.....	63
8. REFERÊNCIAS.....	64

LISTA DE FIGURAS

Figura 1. Estrutura cristal de AtSerp1 de <i>Arabidopsis thaliana</i>	24
Figura 2. Estruturas de serpinas.....	25
Figura 3. Representação da formação do complexo ternário de Michaelis do pentassacarídeo de heparina com a antitrombina.....	27
Figura 4. Alinhamento de TcSERPIN e outras sequências homólogas.....	40
Figura 5. Dendrograma do alinhamento de TcSERPIN com sequências homólogas.....	41
Figura 6. Alinhamento do RCL de serpinas de plantas.....	42
Figura 7: Estrutura tridimensional de TcSERPIN a partir do molde AtSerp1 de <i>Arabidopsis thaliana</i>	43
Figura 8. Estrutura tridimensional das cisteíno-proteases do cacau.....	44
Figura 9: Expressão e purificação de TcSERPIN recombinante.....	45
Figura 10. Avaliação dos espectros da estrutura secundária de TcSERPIN....	46
Figura 11. <i>Unfolding</i> e <i>Refolding</i> de TcSERPIN em pH 7.0 e 5.4.....	47
Figura 12: Teste de atividade de TcSERPIN contra papaína do mamão e tripsina suína.....	48
Figura 13. <i>Docking</i> entre TcSERPIN e as serino- e cisteíno-proteases.....	49
Figura 14. Cluster 1 mostrando as principais proteínas das vias relacionadas a TcSERPIN.....	51
Figura 15. <i>Docking</i> entre a TcSERPIN e as papaínas RD19 e RD21 de <i>T. cacao</i>	52
Figura 16. <i>Docking</i> entre TcSERPIN e a heparina de sódio.....	53
Figura 17. Análise em zimograma atividade de protease capturada de extratos de folhas de <i>T. cacao</i>	55

1. INTRODUÇÃO

Inibidores de proteases do tipo serpinas estão classificados como uma superfamília de inibidores que vêm sendo identificados em todos os seres vivos (RAWLINGS, et al, 2004). Até aqui, milhares de proteínas referentes à superfamília têm sido identificadas, desde mamíferos, vertebrados, invertebrados, plantas, vírus, bactérias e archaea (GETTINS & OLSON, 2016).

Esses inibidores irreversíveis são caracterizados como inibidores de serino-peptidases, de onde advém o nome “serpin” (***Serine proteinase inhibitor***) (CARRELL & TRAVIS, 1985; GETTINS & OLSON, 2016).

As serpinas são proteínas peculiares e nos meios celulares variações na sua estrutura implicam no seu estado funcional. Em sua conformação nos meios celulares seu estado pode alternar entre o nativo, latente, clivado ou polimérico. No estado nativo, o RCL (alça do centro de reação) está exposto e pronto para a interação com a peptidase-alvo. No estado latente, o RCL encontra-se espontaneamente inserido entre as folhas- β A, sendo essa forma instável para as serpinas e que inativa sua ação inibitória. No estado clivado, o RCL, após reação com a protease, se encontra inserido entre as folhas- β A. No estado polimérico, em que, por instabilidades do meio como variações de temperatura, por exemplo, faz com que o RCL seja inserido nas folhas- β A de uma outra serpina, o que possibilita a formação de aglomerados (IRVING et al., 2000, 2002; GETTINS, 2002; HEIT et al., 2013).

O que torna o padrão peculiar das serpinas é o seu modo de inibição. Serpinas se comportam como inibidores do substrato suicida (WHISSTOCK & BOTTOMLEY, 2006). Em que o loop reativo (RCL) age como substrato para a protease. Nesse ponto, o principal sítio ativo da protease reconhece o RCL da serpina como um potencial substrato, isso, no entanto, se configura como uma armadilha para a enzima. No primeiro momento da interação ocorre a formação de intermediário acil-enzima, tal como descrito para um complexo do tipo Michaelis, e ocorre a inserção dos resíduos do RCL, onde ocorreu a clivagem, entre as folhas- β A da serpina. Esse é um passo crucial para serpinas, pois a inserção do loop precisa ser rápido o suficiente para que a peptidase não complete o mecanismo de desacilação. Se bem sucedido, o RCL ao adentrar entre as folhas- β A, modifica sua

conformação, distorcendo seu sítio ativo e arrasta a protease em 70 Å (HUNTINGTON et al., 2000). Como resultado dessa ação, ocorre a formação de um complexo estável e ligado covalentemente. Se, no entanto, a ação da serpina não for bem sucedida, a protease pode completar o mecanismo de desacilação e ser ativada, e a serpina terminar como um substrato, com o RCL clivado (KHAN et al., 2011; GETTINS, 2002). O RCL, principal sítio ativo das serpinas, está posicionado próximo a região carboxi-terminal e é extremamente importante para função inibitória desses inibidores. A região em si, possui uma média de 25 resíduos de aminoácidos altamente conservados (HOPKINGS et al., 1993; IRVING et al., 2000; GETTINS, 2002).

As serpinas mais conhecidas e estudadas são a antitrombina e α_1 -antitripsina, que são proteínas regulatórias que controlam a atividade proteolítica de serino-proteases da família da quimotripsina e tripsina (SILVERMAN et al., 2001; GETTINS et al., 2002). Embora sejam reconhecidas por serem inibidores de serino-proteases, também são caracterizadas por sua capacidade de inibição a cisteíno-proteases (OTLEWSKI et al., 2005). Nem todas as serpinas possui capacidade inibitória, pois, evolutivamente, algumas serpinas parecem ter perdido essa capacidade. Membros como ovoalbumina e maspin, por exemplo, pertencem a essa categoria (GETTINS, 2002).

As serpinas estão envolvidas em diversos mecanismos celulares de regulação e homeostase (GETTINS, 2002), e têm sido relacionadas a supressão e desenvolvimento tumoral (WHISSTOCK & BIRD, 2017; ROBERTS & HEJGAARD, 2008), mecanismos apoptóticos (WILLIAMS & DICKMAN, 2008), cascatas de coagulação, processos inflamatórios e regulação do sistema do complemento (HEIT et al., 2013).

Nos bancos de dados, estudos com serpinas visando estratégias de desordem em humanos são diversos (LAW et al., 2006; CHEN et al., 2015; LUENGO-GI et al., 2016; WHISSTOCK & BIRD, 2017; LUCAS et al., 2018). No entanto, o estudo desses inibidores para melhor conhecimento em plantas, embora venha aumentando, ainda não alcança os diversos estudos com serpinas em humanos. No banco de dados do PDB, por exemplo, diversas estruturas cristalizadas de serpinas de humanos e invertebrados têm sido relatadas contra somente uma de plantas, a AtSerp1 de *A. thaliana* (LAMPL et al., 2010).

Entre os representantes de serpinas de plantas, podemos citar a proteína Z da cevada (ROBERTS & HEJGAARD, 2008), a primeira com estudos relacionados à expressão, purificação e caracterização (DAHL et al. 1996a; RASMUSSEN et al. 1996), e, até os dias atuais, permanece como sendo uma das mais estudadas (LI et al., 2014). Outros representantes bem estudados também são as serpinas do trigo, centeio, aveia, algodão, pepino, tomate, abóbora e *Arabidopsis* (ROBERTS & HEJGAARD, 2008). No entanto, nenhum dos trabalhos esclarece as rotas nas quais as serpinas, de fato, estão inseridas na planta (LAMPL et al., 2013).

Estudos *in vitro* com AtSerp1 têm caracterizado essa proteína com potencial inibitório contra uma metacaspase 9 (AtMC9) (VERCAMMEN et al., 2006) e a metacaspase 1 (AtMC1) (LEMA ASQUI et al., 2017) e uma papaína, a RD21 (LAMPL et al., 2010). Esses trabalhos sugerem que esses inibidores possam estar relacionados com vias de defesa da planta (ALVAREZ-ALFAGEME et al., 2011) e, provavelmente, morte celular programada (SCHALLER, 2004; LI et al., 2008). Os trabalhos mais recentes com a AtMC1 demonstram que a AtSerp1 possa estar envolvida não somente nos processos de defesa e morte celular, mas também em processos catalíticos da planta (LEMA ASQUI et al., 2017).

O desenvolvimento de vias de morte celular programada é fundamental na resposta de plantas em mecanismos de estresse e defesa contra patógenos. Num contexto mais amplo, os sistemas de defesa do patógeno e do hospedeiro são acionados no momento da interação. E, nessa linha de defesa atuam algumas cisteíno-proteases (PLCPs) no apoplasto da planta, ou, no citoplasma dos invasores (SHINDO & VAN DER HOORN, 2008).

As PLCPs atuam protegendo a planta da invasão contra patógenos. Contudo, a história não é tão simples, porque o invasor também desenvolve mecanismos de defesa e ação sintetizando efetores (fungos), outras PLCPs e inibidores para vencer a barreira de defesa da planta. Nesse contexto, não somente as cisteíno-proteases da planta atuam, mas também cisteíno-proteases dos patógenos. Essas interações geram um verdadeiro “campo de batalha”, onde proteases e inibidores atuam como atores em seus papéis de defesa e ataque tanto da planta como dos seus patógenos nos espaços intercelulares (SHINDO & VAN DER HOORN, 2008; HOU et al., 2018).

Devido a interação que ocorre entre planta-patógeno com relação a proteínas e enzimas envolvidas durante os mecanismos de ataque e defesa

(SHINDO & VAN DER HOORN, 2008) estudos moleculares em plantas precisam ser dirigidos para o entendimento dessa relação durante a infecção. Sendo assim, o presente trabalho está relacionado com enzimas e inibidores de *Theobroma cacao*.

Theobroma cacao é planta perene originária da floresta tropical úmida na Amazônia (GRIFFITH et al., 2003). O gênero *Theobroma* pertence à família Malvaceae (LORENZI & SOUZA, 2012), e ganhou destaque mundial na economia graças ao sabor peculiar advindo de sua amêndoas (JUDD et al., 2009), que são amplamente aproveitadas na produção de chocolate, bebidas, geléias, sucos, alguns produtos cosméticos, etc (LORENZI & MATOS, 2002; GESTEIRA et al., 2007). Contudo, a susceptibilidade do cacau ao ataque de patógenos fúngicos e insetos tem afetado a estabilidade da produção em escala mundial. Ataques de insetos como *Monalonian* spp. na América Latina, *Conopomorpha cramerella* na Ásia e besouros do gênero *Xyleborus*, principal vetor de *Ceratocystis cacaofunesta*, um fungo causador da murcha de Ceratocystis estão entre os exemplos de pragas que assolam o cacaueiro (PLOETZ, 2007). Ainda, entre outros fungos que atacam o cacaueiro e causam doenças estão fungos do gênero *Moniliophthora*, que causam injúrias como a monília e a vassoura-de-bruxa (AIME & PHILLIPS-MORA, 2005; EVANS et al., 2013). Doenças causadas por esses patógenos podem causar perdas de até 20% na produção total das amêndoas (PLOETZ, 2007).

Devido a susceptibilidade do cacaueiro a patógenos como insetos e fungos, é que estudos visando o entendimento das relações entre planta e patógeno têm sido desenvolvidos. Nessas relações ocorrem constante acúmulos e degradação de proteínas, sejam enzimas, inibidores ou efetores (SHINDO et al., 2008). Informações sobre os mecanismos bioquímicos e moleculares implicados durante a interação de *T. cacao* e seus patógenos têm sido necessários na busca de melhorias, visando o conhecimento e a busca por respostas (Gesteira et al., 2007; Mondego et al., 2008; Argout et al., 2008, Argout et al., 2011; Barbosa et al., 2018). Dessa forma, o conhecimento dessas vias no cacau e em seus patógenos durante a interação são necessários. Sendo assim, o presente trabalho tem por objetivo caracterizar uma serpina de *Theobroma cacao*, determinar sua função e gerar diretrizes que possam ajudar a planta ao combate de patógenos.

2. OBJETIVOS

2.1.OBJETIVO GERAL

Caracterizar *in silico* e *in vitro* uma serpina de *Theobroma cacao*.

2.2. OBJETIVOS ESPECÍFICOS

- i. Caracterizar estruturalmente e analisar a interação com possíveis alvos *in silico*;
- ii. Produzir a proteína por vias heterólogas e analisar a estrutura secundária e estabilidade térmica;
- iii. Avaliar a interação da proteína com alvos *in vitro*.

3. REFERENCIAL TEÓRICO

3.1. Proteases

O termo protease muito utilizado na literatura é sinônimo de peptidase. As peptidases pertencem a classe das hidrolases e são enzimas proteolíticas (RAWLING & BARRET, 1999) que causa a hidrólise das ligações peptídicas, uma reação em que ocorre a adição de uma molécula de água (LEHNINGER et al., 2000).

Desde 1984, a utilização do termo “peptidase” tem sido incentivado pela União Internacional de Bioquímica e Biologia Molecular (UIBBM) e pelo banco de dados público do MEROPS para o subgrupo das hidrolases (SANTOS et al., 2016). De acordo com a nomenclatura enzimática (<http://www.sbcs.qmul.ac.uk/iubmb/>), as peptidases pertencem à classe enzimática (do inglês, *enzyme class* – EC) 3 (hidrolases), subclasse 3.4 (peptidases) (TIPTON, 1992; SANTOS et al., 2016). E, de acordo com o MEROPS estão divididas em 62 clãs e 268 famílias, segundo evidências evolutivas, similaridades de sequências e estruturas terciárias (RAWLINGS et al., 2018).

As peptidases também são divididas de acordo com as reações que catalisam, sendo compreendidas em dois grupos: as exopeptidases, enzimas que clivam ligações peptídicas de resíduos de aminoácidos em regiões N-terminais (aminopeptidases) ou C-terminais (carboxipeptidases) e as endopeptidases (proteínases), enzimas que realizam a clivagem das ligações peptídicas no interior da cadeia polipeptídica (RAWLINGS & BARRETT, 1993; SANTOS et al., 2016).

Além da classificação pelo modo de clivagem na cadeia polipeptídica, essas enzimas proteolíticas também são classificadas de acordo com seus mecanismos catalíticos relacionados ao seu sítio ativo (RAWLINGS & BARRETT, 1993), as relações evolutivas baseadas na sequência primária de seus aminoácidos e sua estrutura terciária (SANTOS et al., 2016). E, de acordo com esse critério, levando em consideração seus mecanismos catalíticos, as carboxipeptidases e as endopeptidases estão divididas em sete subclasses: serino, cisteíno, aspartil, metalo (RAWLINGS & BARRETT, 1993), treonino (SEEMÜLLER et al, 1995), glutamato (FUJINAGA et al., 2004) e mais recentemente, asparagino (RAWLINGS et al., 2011).

As classes conhecidas serina, cisteína, aspartil e metalo podem ser caracterizadas da seguinte forma: as peptidases do tipo serino, possui uma serina no centro do seu sítio ativo, sendo a serina o principal resíduo de aminoácido catalítico; as cisteíno-proteases, um resíduo de cisteína; as aspartil-proteases utilizam dois resíduos de ácido aspártico e; as metaloproteases, um íon metálico (geralmente zinco) (RAWLINGS & BARRETT, 1993).

A quinta classe conhecida contempla as treonino-proteases. Antes, essas peptidases eram conhecidas apenas por serem componentes do proteassoma, a partir de pesquisa descobriu-se que, das 14 subunidades componentes do proteassoma, três eram peptidases e possuíam uma treonina em seu sítio ativo, próximo a região N-terminal (SEEMULLER et al., 1995).

As endopeptidases do tipo glutâmico foram encontradas em fungos (equilisinias) e corresponde a sexta classe encontrada (FUJINAGA et al., 2004). A sétima e última classe localizada até os dias atuais é a das asparagino-proteases, com uma asparagina central no seu sítio ativo (RAWLINGS et al., 2011).

A importância das proteases deve-se ao fato de que elas estão envolvidas em vários processos biológicos importantes à célula, como controle, regulação, diferenciação e morte celular, e estão presentes em muitas vias de sinalização e rotas metabólicas de plantas e animais. Dentre outras funções podem também realizar clivagem na cadeia proteolítica de outras enzimas e ativar precursores inativos (zimógenos) para que estes possam se tornar ativos e exerçam sua função na célula (RYAN, 1990; BERG, LEHNINGER et al., 2000).

Um outro papel importante que as proteases desempenham na célula, é a degradação e reciclagem de proteínas, disponibilizando aminoácidos que podem ser reciclados e reutilizados pelo organismo. Um excelente exemplo desse mecanismo a ser citado, está no processo digestório, onde as proteínas ingeridas na alimentação são degradadas em pequenos peptídeos e aminoácidos, e, assim, são absorvidas pelo organismo. Enzimas como a pepsina, tripsina e a quimotripsina, exercem essa função de quebra das proteínas e são fundamentais durante os processos de digestão e absorção. De maneira geral, reações que envolvem proteases são importantes não somente na quebra de outras proteínas, mas também para a regulação de muitas enzimas e proteínas (BERG, TYMOCKZO E STRYER, 2008).

3.2. Mecanismo catalítico das proteases

A função catalítica que envolve as proteases, de maneira geral, está baseada em um mecanismo de ação por meio de uma tríade catalítica. Uma região conservada que possui três resíduos de aminoácidos que interagem entre si. Nessa interação um resíduo central realiza um ataque nucleofílico a uma região específica do substrato, e essa interação é suportada pela estabilidade de ligações de hidrogênio entre os três resíduos da tríade. Seguindo o exemplo do mecanismo da quimotripsina, uma serino-protease, que possui na sua região de sítio ativo os seguintes aminoácidos compondo a tríade: serina (Ser), aspartato (Asp) e histidina (His). Nesse mecanismo de interação, a tríade é suportada pela serina 195 ligada por interações de hidrogênio com o anel imidazólico da histidina 57, sendo que o grupamento NH do anel imidazólico está ligado também ao carboxilato do aspartato 102. A serina central da tríade é altamente reativa e realiza um ataque nucleofílico levando a clivagem das proteínas. Essa interação ocorre mediante uma hidrólise que, geralmente, acontece em duas etapas: acilação e desacilação (LEHNINGER et al., 2000; KRAUT, 1977; BULLER & TOWNSEND, 2013).

Durante a fase de acilação, o oxigênio da serina 195 realiza um ataque nucleofílico ao carbono da carbonila da ligação peptídica formando um estado de interação tetraédrico instável com quatro átomos ligados ao carbono carbonílico. O tetraedro formado é instável e possui uma carga negativa no átomo de oxigênio derivado da carbonila, nesse passo, a histidina 57 ligada por ligações de hidrogênio com a serina 195, age como uma base geral e retira um próton da serina, e o aspartato 102 carregado negativamente, estabiliza a carga positiva em torno da histidina. No passo seguinte, o intermediário tetraédrico colapsa e dá origem a uma acil-enzima. Nessa fase, a histidina 57, pode atuar como um doador de prótons para a amina derivada da clivagem da ligação peptídica que está livre para sair da enzima, completando a fase de acilação (LEHNINGER et al., 2000; KRAUT, 1977).

Durante a fase de desacilação a ligação éster formada na etapa anterior é hidrolisada e ocorre a regeneração da enzima. Nessa etapa, uma molécula de água entra e toma o lugar que antes era ocupado pela amina do substrato. Nesse passo, a histidina 57 funciona como um acceptor de prótons e retira um próton da molécula de água, o que leva novamente ao estado de transição tetraédrico, que, posteriormente, colapsa em um ácido carboxílico e a reação termina com a ativação

da enzima livre e substrato clivado (LEHNINGER et al., 2000; KRAUT, 1977). A velocidade em que as etapas de acilação e desacilação ocorrem depende da variação em relação ao substrato utilizado em cada reação (BULLER & TOWNSEND, 2013).

O mecanismo relatado anteriormente tem por base reações catalíticas de proteases do tipo serina. No entanto, ainda que as trajetórias das reações possuam diferença na posição dos átomos, nos mecanismos de transferência de prótons dos resíduos envolvidos e na composição do tipo da tríade envolvida, todas as peptidases convergiram evolutivamente para reações que ocorrem a partir de uma fenda (buraco do oxianion). A estereoquímica da catálise dessas enzimas são extremamente conservadas em todas as peptidases, e, embora os grupos do sítio ativos das proteases exijam diferentes geometrias de transferência de prótons, a depender do tipo de protease, o mecanismo de ligação ao substrato permanece conservado (BULLER & TOWNSEND, 2013).

3.3. Inibidores protéicos de proteases

Devido a importância das proteases nos mecanismos celulares, estas moléculas precisam ser reguladas, pois uma peptidase fora de controle dentro da célula pode causar doenças, e, até mesmo a morte. Uma maneira de regulação celular são as formas inativas (zimógenos) dessas peptidases, outra maneira é por intermédio dos inibidores (RYAN, 1990).

De modo geral, os inibidores possuem a capacidade de inibir a ação de outras enzimas, e isso faz dessas proteínas ferramentas fundamentais no controle e regulação de diversos mecanismos celulares e de defesa da célula (HAQ et al., 2004).

No banco de dados do MEROPS são encontradas informações relevantes sobre os vários inibidores de peptidases e suas respectivas classificações. Esses inibidores podem ser classificados como protéicos, sendo assim divididos em famílias e clãs de acordo com a sequência primária dos seus resíduos de aminoácidos e estrutura tridimensional; ou, ainda, como pequenas moléculas, importantes como reagentes e drogas em diversas pesquisas. Nesse último caso, no entanto, não há uma classificação específica, mas um índice separado é fornecido sobre eles (RAWLINGS et al., 2018).

De maneira geral, o banco de dados dispõe de informações importantes sobre os inibidores protéicos, tais como nome, estrutura química, proteases-alvo, mecanismos inibitórios, histórico, referências bibliográficas, alinhamentos das sequências, árvore filogenética e a classe à qual pertencem, entre outras informações (SANTOS et al., 2016; RAWLINGS et al., 2018). Da mesma maneira que as peptidases, esses inibidores estão organizados em famílias e clãs. Até 2015, a tabela do MEROPS apresentava uma divisão em 39 clãs e 79 famílias. Atualmente, uma nova divisão apresenta um aumento para 82 famílias de inibidores, todas seguidas por uma letra e um número que as identificam, e o tipo de inibição correspondente (RAWLINGS et al., 2018).

Devido a importância que as peptidases têm em diversos processos celulares de regulação, o controle dessas enzimas se faz necessário (CRAIK et al., 2011). Sendo assim, o controle dessas peptidases mediante a uma inibição seletiva tem se tornado importante frente a diversas pesquisas visando o conhecimento e o tratamento de diversos processos inflamatórios, imunológicos e respiratórios, alguns tipos de infecções, além de doenças cardiovasculares e degenerativas como a doença de Alzheimer, por exemplo, que atualmente está em fase de testes clínicos com compostos inibitórios da β -secretase (do inglês *beta-amyloid-converting enzyme* – BACE) (SANTOS et al., 2016).

Aproveitando-se do potencial biotecnológico desses inibidores muito têm sido os esforços para o conhecimento dessas proteínas no controle de proteases desreguladas relacionadas com diversas doenças, tais como asma, artrite reumatóide, trombose e algumas doenças inflamatórias (ABBENANTE & FAIRLIE, 2005).

Os mecanismos inibitórios podem ocorrer de três formas: por interações do tipo canônica, não canônica e de serpinas (KROWARSCH et al., 2003). No tipo de interação canônica, o inibidor interage com sítio ativo da enzima de maneira semelhante ao que ocorre com a enzima e o substrato, levando a formação de um complexo bastante estável e de difícil dissociação, é uma inibição do tipo competitiva. Já no tipo de interação não canônica, o inibidor interage com porções N ou C-terminais da enzima e não há competição com o substrato. Os inibidores do tipo serpinas, no entanto, reagem com a enzima de maneira similar a interação canônica. Porém, nesse tipo de interação, diferente da canônica, a ligação é instável (LOEBERMANN et al., 1984), o que leva a clivagem da alça (RCL) da serpina e

deslocamento da protease-alvo a um ângulo de 70°. Sendo que, na etapa final da reação, a protease estará presa a serpina por meio de uma ligação covalente, irreversível (GETTINS, 2002).

3.3.1. Serpinas

As serpinas são uma superfamília de inibidores de peptidases encontrada em todos os seres vivos e até mesmo em vírus, com exceção do reino fungi. (RAWLINGS, et al, 2004). O nome da superfamília se deve ao fato de sua inibição, primeiramente, ter sido reconhecida nas serino-proteases, do qual derivou o nome atualmente conhecido como serpinas (*Serine Protease Inhibitor*) (HUNT & DAYHOFF, 1980).

Entre os primeiros e mais bem estudados representantes dessa superfamília estão a antitrombina, α 1-antitripsina e ovoalbumina em humanos. Desde o início dos estudos com essas proteínas, foi observado que esse grupo de inibidores possuía sequências conservadas em determinadas regiões, principalmente numa região de loop específico por onde ocorriam as interações com suas proteases-alvo. A estrutura tridimensional dessas proteínas demonstra-se extremamente conservadas e semelhantes em todas as espécies (GETTINS, 2002).

O MEROPS posiciona as serpinas no clã ID, família I4 que contém os inibidores do tipo serino- e cisteíno-proteases (RAWLINGS et al., 2018). Estão divididas em 16 clados, indo de A a P e são organizados da seguinte maneira: A (inibidor de protease α -1), B (ovoalbumina intracelulares), D (cofator de heparina II), E (nexina), F (α -1-antiplasmina), G (inibidores de C1), H (HSP41), I (neuroserpina), J (serpinas do caranguejo ferradura), K (serpinas de insetos), L (serpinas de nemátodos), M (serpinas de *Schistosoma*), N (serpinas do vírus SPI1-2/CrmA-like), O (serpinas do vírus SPI3) e P (serpinas de plantas) (IRVING et al., 2000; SILVERMAN et al., 2001).

Essas proteínas são caracterizadas por serem inibidores irreversíveis de serino-proteases pertencentes à família da tripsina e quimotripsina e cisteíno-proteases (OTLEWSKI et al., 2005). Nem todas as serpinas tem capacidade inibitória, pois, aparentemente, algumas dessas serpinas, como a ovoalbumina, por exemplo, parecem ter perdido sua capacidade inibitória no decorrer do histórico evolutivo. De maneira geral, são proteínas versáteis e atuam em diversos

mecanismos celulares desde o transporte de hormônios; como chaperona; proteínas de choque-térmico; armazenamento (ovoalbumina) (GETTINS, 2000; RUSSO, 2006); na supressão de genes tumorais (Maspin) (ROBERTS & HEJGAARD, 2008) e em outros mecanismos relacionados a controle, regulação e apoptose celular (WILLIAMS & DICKMAN, 2008). Estão também presentes no controle de vias proteolíticas intra e extracelular. Entre outros exemplos a serem citados estão a antitrombina, uma serpina encontrada no plasma humano, que age no controle da coagulação sanguínea e o inibidor C1 que controla a ativação do complemento (ELLIOTT et al., 1998; HUNTINGTON et al., 2000). Outra serpina humana, α 1-antitripsina está relacionada a processos inflamatórios e supressão de genes tumorais (RAWLINGS et al., 2004), sendo esse último, semelhante a Maspin, com relação aos processos de supressão a desenvolvimento tumorais (ROBERTS & HEJGAARD, 2008). Outra representante, a serpina B6 de humanos, possui potencial inibitório frente a catepsina G (componente de grânulos azurófilos), sendo, então, relacionada à regulação do sistema imunológico (HEIT et al., 2013).

Ainda, muitas serpinas têm demonstrado real importância em diversos estudos, dentre alguns podemos citar representantes como a Maspin, uma serpina humana que está relacionada na supressão e desenvolvimento da metástase em células cancerígenas da mama e próstata (DZINIC et al., 2017); serp-1, do vírus Myxoma do coelho, contra doenças arteriais coronarianas, já em fase de testes clínicos (CHEN et al., 2013, 2015); uma serpina de arroz transgênico, para estudos relacionados ao controle da pressão arterial (YANG et al., 2006); e AtSerp1 de *Arabidopsis thaliana* como potencial inibidor de proteases do sistema digestivo de insetos (ALVAREZ-ALFAGEME et al., 2011), entre outros.

3.3.1.1. Mecanismos estruturais funcionais das serpinas

Em sua estrutura, serpinas demonstram um padrão geral de conservação estrutural, sendo constituídas por 3 feixes de folhas β (5 filamentos s A, 6 filamentos s B e 4 filamentos s C), entre 7 a 9 α -hélices (hA-hI) e uma alça do centro de reação (RCL – do inglês, *reactive centre loop*) (HOPKINGS, et al., 1993; IRVING et al., 2000). O destaque deve ser dado a esse último, o RCL, pois é o principal sítio ativo do inibidor, e é o local onde ocorre a interação entre a serpina e a protease.

O RCL possui em média 25 resíduos de aminoácidos e apresentam um padrão de conservação desses resíduos muito importante para a atuação de sua atividade inibitória (OTLEWSKI et al., 2005; GETTINS, 2000).

Em sua configuração estrutural, o RCL possui determinadas regiões que são fundamentais para a sua estabilização, entre elas estão: a “*hinge region*” (região de dobradiça) localizada, em média, entre a posição P15 a P9, com 6 resíduos de aminoácidos em média, e que é importante para que haja maior mobilidade desses inibidores entre as mudanças conformacionais sofridas no momento da interação com a protease; a “*breach region*”, local onde o RCL é inserido após a clivagem resultante de sua interação com a enzima-alvo; a região “*shutter domain*”, que ao lado da região “*breach*”, facilita a abertura das folhas- β para que a região de dobradiça do RCL possa ser inserida entre as folhas- β ; e a região “*gate*”, fundamental durante a inserção do RCL dentro das folhas- β (IRVING et al., 2000; GETTINS, 2002) (Figura 1).

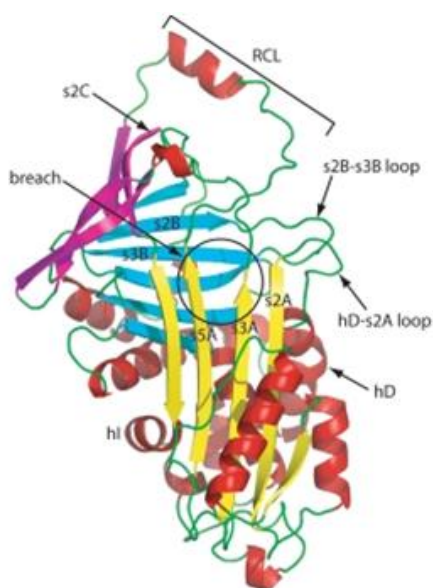


Figura 1. Estrutura de AtSerpin1 de *Arabidopsis thaliana* mostrando partes conservadas de uma serpina. A estrutura foi resolvida por cristalografia e difração de raio-X (LAMPL et al., 2010).

Os mecanismos inibitórios próprios das serpinas envolvem mudanças conformacionais importantes em sua estrutura para que haja sucesso na inibição da protease-alvo. De maneira similar ao que ocorre entre as proteases e o substrato, a protease-alvo interage entre a posição P1 e P1' do RCL. Essa interação leva a formação de um complexo não covalente do tipo-Michaels, e assim como ocorre

com o substrato, a alça do centro de reação é clivada levando a formação de um estado de transição tetraédrico. A catálise segue até a formação do último intermediário acil-enzima, e, a partir daí, dois caminhos podem ser seguidos, no primeiro, o intermediário acil-enzima é hidrolisado e a serpin é liberada com o RCL clivado e a enzima livre para outra catálise. E no segundo caminho, o RCL após a clivagem é inserido dentro da folha- β s A, levando a formação de um filamento extra, o que gera uma mudança conformacional no interior da serpin e o deslocamento da protease da região de interação do RCL para o lado oposto (em torno de 70Å) da serpin, levando a formação de um complexo covalente e inibindo irreversivelmente a protease (Figura 2) (GETTINGS, 2002).

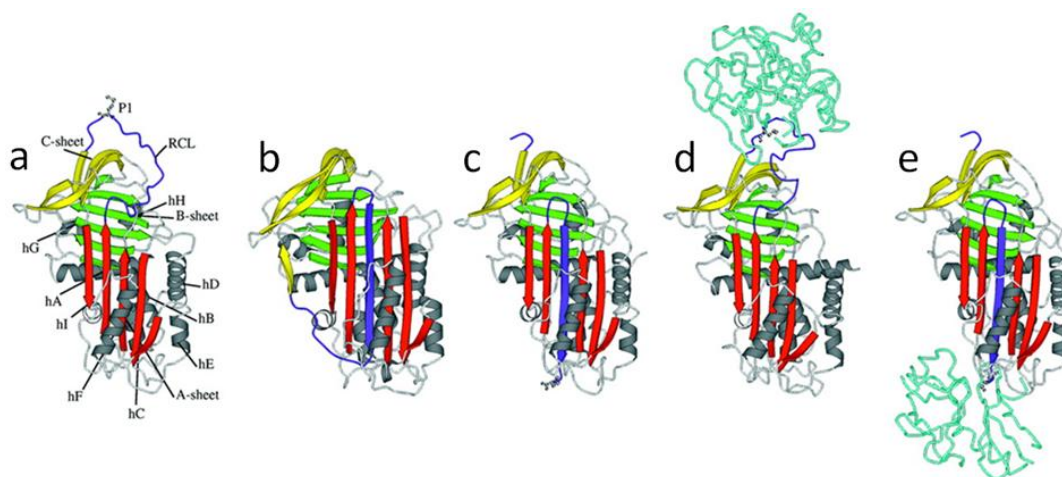


Figura 2. Estruturas das serpinas: **a**, α -1 AT nativa (PDB 1QLP); **b**, ATIII latente (PDB 2ANT); **c**, α -1 AT clivado (7API PDB); **d**, complexo de Michaelis entre Serpin 1 (Alaserpin de *Manduca sexta*) e tripsina (PDB 1I99); **e**, complexo covalente entre α -1 AT e tripsina (PDB 1ezx) (SILVERMAN et al., 2001).

3.3.2. Serpinas e a afinidade com pentassacarídeos

Além da versatilidade de ligação a diversas proteases, seja do tipo serina ou cisteína, as serpinas também tem a incrível capacidade de se ligar a outros ligantes não-protéicos. Um exemplo disso é a afinidade entre serpinas e heparina.

Entre as serpinas com afinidade à heparina pode-se citar a antitrombina, o cofator II de heparina, o inibidor da proteína C, o inibidor do ativador do plasminogênio-1, nexin-1 e kallistati (PATSTON et al., 2004). Dentre esses, o representante mais conhecido por ter uma enorme afinidade com a heparina é a

antitrombina. Esse inibidor é muito presente no sangue e regula a ação de proteinases que participam das cascatas de coagulação sanguínea, porém, sua ação inibitória se dá de maneira muito lenta quando este não está complexado com a heparina. A interação da antitrombina com ligantes do tipo heparina ou sulfato de heparina fazem com que a ação inibitória da antitrombina seja potencializada e esta se torne uma proteína inibitória com maior velocidade nas taxas de inibição, onde há relatos desse aumento em até 50 vezes (VAN BOVEN & LANE, 1997). Um fato interessante é que, a presença de sulfato de heparina na superfície das células endoteliais dos vasos sanguíneos parecem ativar a antitrombina presente no meio sanguíneo, isso pode explicar o porquê desses vasos terem potencial anti-trombótico (OLSON et al., 2010).

A heparina é um polissacarídeo com peso molecular entre 3000 a 30000 Da. Não somente possui a capacidade de ativar determinadas proteínas como alguns grupos de serpinas, mas também está muito presente na base de medicamentos de origem farmacêutica, pois tem capacidade de prevenir a TVP (Trombose Venosa Profunda) em pacientes com alto risco trombótico após algumas cirurgias, portadores de neoplasias ou sepse. Também é muito utilizada em procedimentos cirúrgicos mais delicados e pode favorecer o tratamento de casos trombóticos mais severos (HAMERSCHLAK et al., 1996).

Os detalhes de como ocorre a interação entre serpina/heparina foram possíveis por meio de estudos com as estruturas cristalografadas em raio-x tanto da antitrombina livre como a forma complexada com a heparina (CARRELL et al., 1994; JIN et al., 1997). Aparentemente, três regiões da antitrombina são importantes nessa interação: a extremidade N-terminal da hélice A, toda a hélice D e a região N-terminal do loop. Nessas regiões, resíduos com potencial positivos como a Lys11, Arg13, Arg46, Arg47, Lys114, Lys125 e Arg129 interagem por ligações de hidrogênio com os grupos sulfato e carboxilato carregados negativamente na heparina (OLSON et al., 2010).

Tanto a heparina como o sulfato de heparina ativam o potencial inibitório da antitrombina por aumentar sua afinidade ao alvo. Os glicosaminoglicanos como a heparina podem ativar o aumento da interação entre inibidor/protease por fazer uma ligação que une tanto a serpina quanto a proteinase, promovendo a formação de um complexo ternário (OLSON & BJÖRK, 1991). Essa ativação do pentassacarídeo possibilita uma reatividade maior da antitrombina pelos fatores Xa e IXa, e também

da trombina (OLSON et al., 2010; GETTINS & OLSON, 2016) (Figura 3). Aparentemente o local onde ocorre a ligação entre o pentassacarídeo e a serpina faz com que a região N-terminal e a hélice A sofram reorientações que aumentam o local de ligação, indo da hélice D e se estendendo por toda a estrutura da serpina (WHISSTOCK et al., 2000). A extensão da hélice D encurta a separação entre a protease e a serpina, e, diversas mudanças conformacionais nas hélices da serpina parecem levar a um aumento da afinidade e, conseqüentemente, maior rapidez nas taxas inibitórias (OLSON et al., 2010).

No entanto, ainda não há relatos de que serpinas vegetais possam interagir com proteases, de modo facilitado, pela formação de complexo ternário com algum outro polissacarídeo existente na planta.

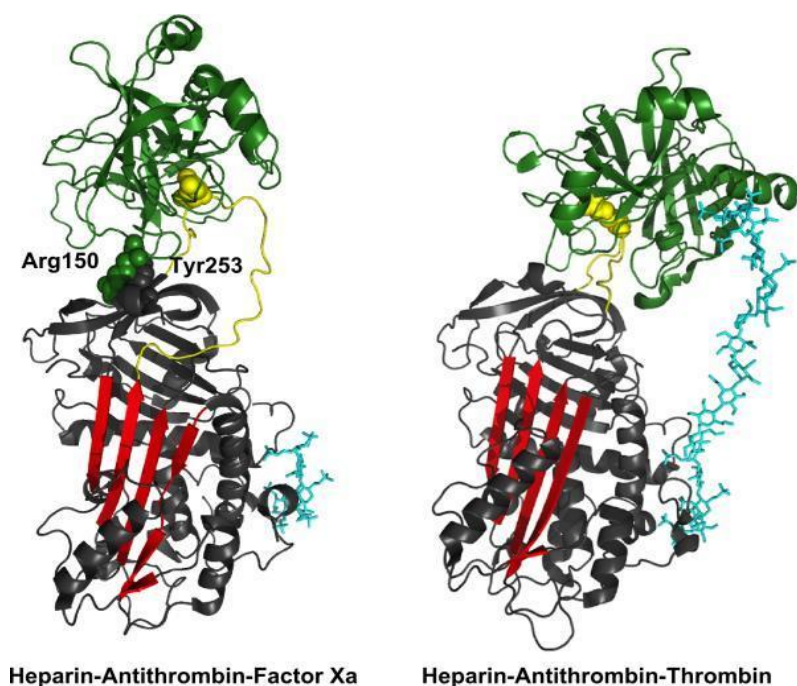


Figura 3. Representação da formação do complexo ternário de Michaelis do pentassacarídeo de heparina com a antitrombina: Imagem à direita (1TB6 PDB); e com fator Xa à esquerda (2GD4 PDB) (OLSON et al., 2010).

3.3.3. Serpinas de plantas

Em humanos, o estudo com serpinas tem sido extensivamente explorado, e muitas serpinas de humanos têm sido caracterizadas como potentes inibidores, tais

como antitripsina envolvida em respostas inflamatórias (LAW et al., 2006), a antitrombina e o co-fator de heparina II envolvidas em cascatas de coagulação e fibrinólise, a Maspin na supressão de tumores, o inibidor C1, relacionado a componentes do complemento (C1r e C1s), entre outras (GETTINGS, 2004). Já em plantas, os estudos com essas serpinas datam desde os anos 70, mas somente nas últimas duas décadas, trabalhos têm sido realizados visando à compreensão dos mecanismos de funcionamento (ROBERTS & HEJGAARD, 2008).

Aparentemente, a primeira serpina de plantas que teve sua sequência elucidada foi uma serpina de cevada, a “proteína Z”, no final da década de 70. Devido a esse fato, no banco de dados público, muitas serpinas aparecem acompanhadas por uma letra “Z” para designar a similaridade com a primeira serpina de planta identificada, “proteína Z” da cevada (ROBERTS & HEJGAARD, 2008). A designação Zx muitas vezes observada nos bancos de dados tem por base a comparação da homologia de algumas serpinas de plantas na região do RCL (ROBERTS & HEJGAARD, 2008).

As serpinas identificadas em cevada e trigo estão entre as primeiras serpinas recombinantes de plantas expressas e purificadas (DAHL et al. 1996a; RASMUSSEN et al. 1996). O grupo com maior número de estudos, aparentemente, é o das monocotiledôneas cevada, trigo, centeio e aveia (HEJGAARD, 1976; ROBERTS & HEJGAARD, 2008). Outras plantas também estudadas têm sido o algodão, o pepino, o tomate, a abóbora e *Arabidopsis* (ROBERTS & HEJGAARD, 2008).

Entre alguns estudos realizados com serpinas em plantas, podemos citar a expressão de cinco serpinas no trigo com potencial inibitório contra diversas proteases (OSTERGAARD et al., 2000) e o estudo com duas serpinas da cevada cuja capacidade inibitória frente a tripsina, quimotripsina e catepsina G foi demonstrada, sugerindo função regulatória dentro da planta (DAHL et al., 1996a). No entanto, uma dessas proteínas da cevada, a BSZx, demonstrou ter potencial biotecnológico para proteases humanas ao interagir com alguns fatores de coagulação e formar complexos estáveis com a calcicreína plasmática, o fator VIIa/STF, o fator Xa e com a trombina. O estudo, ainda mostrou que a BSZx atingiu taxas de inibição semelhantes a da antitrombina III de humanos, de forma independente de heparina (DAHL et al., 1996b).

Em *Arabidopsis thaliana*, Vercammen e colaboradores (2006) identificaram uma serpina (AtSerp1) como um inibidor em potencial da metacaspase 9 (AtMC9), cuja função ainda não foi elucidada (VERCAMMEN et al., 2006). Posteriormente, Lampl e colaboradores (2010) cristalografaram AtSerp1 de *Arabidopsis*, sendo que esta é a primeira e única serpina de planta com estrutura cristal até o momento. Testes *in vitro* de AtSerp1, ainda no mesmo trabalho, avaliaram seu potencial inibitório contra uma cisteíno-protease (RD21) que pode estar relacionada em vias de estresse e morte celular da planta. E, em conclusão, os pesquisadores constataram a existência de interação *in vitro* de AtSerp1 e a RD21 (LAMPL et al., 2010). Outros estudos potenciais com AtSerp1 também vêm demonstrando que esta serpina tem potencial inibitório também frente a proteases do sistema digestivo de insetos (ALVAREZ-ALFAGEME et al., 2011).

Em se tratando de sua localização nos organismos, não há um consenso até o momento. Alguns estudos apontam a presença de serpinas no cloroplasto do arroz (TANAKA et al. 2004), no floema da abóbora (YOO et al., 2000), em sementes de maçã e cevada (ROBERTS & HEJGAARD, 2008), *Arabidopsis* (AtSerp1) com expressão diferencial em raízes, folhas e flores, além da confirmação de sua presença no apoplasto (LAMPL et al., 2010; VERCAMMEN et al., 2006).

De modo geral, não há clareza das vias das serpinas em plantas, e, embora estas tenham demonstrado inibição potente e irreversível de proteases *in vitro*, perguntas sobre sua atuação dentro dos organismos vegetais não parecem claras. O que se sabe, até o momento, é que serpinas de plantas possuem capacidade inibitória frente a proteases do tipo serina e/ou cisteína, no entanto, outros estudos devem ser realizados e explorados para uma caracterização segura da via desses inibidores nos vegetais.

3.4. Serpinas em vias de morte celular programada

Patógenos como fungos, por exemplo, secretam efetores no apoplasto e no citoplasma da planta para driblar a linha de defesa imposta (TORUÑO et al., 2016). O parasita ao invadir os tecidos da planta busca desestruturar o mecanismo de defesa do hospedeiro e degradar sua linha de defesa, por meio de enzimas e inibidores. Durante essa interação é travada uma “batalha”, onde proteínas e enzimas são constantemente degradadas tanto pelo patógeno como pelo

hospedeiro. Entre as linhas de defesa da planta e do patógeno destacam-se as cisteíno-proteases do tipo papaína (PLCPs) e diversos inibidores que atuam nessa interação (SHINDO et al., 2008).

Estudos sobre morte celular programada têm sido realizados visando compreender as interações que ocorrem quando um patógeno infecta a planta ou quando essa é submetida a estresse severo. O que se sabe é que determinadas proteases desempenham um papel fundamental na resposta de defesa. Porém, ainda não são compreendidos muito bem os mecanismos envolvidos. Nesse aspecto, o estudo de inibidores, como as serpinas, por exemplo, têm sido implicados como reguladores que restringem a atividade de determinadas proteases que podem estar relacionadas a essas vias (LAMPL et al., 2013).

Estudos com AtSerp1, por exemplo, têm sido desenvolvidos numa tentativa de entender a função das serpinas em plantas e seu possível papel nos mecanismos associados à morte celular programada em contexto de estresse severo ou invasão por patógenos (VERCAMMEN et al., 2006; LAMPL et al., 2010, 2013). Em 2010, e, posteriormente, em 2013, Lampl e colaboradores demonstraram que uma serpina de *Arabidopsis thaliana* possui a capacidade de inibir uma protease do tipo papaína, a RD21, durante o mecanismo de morte celular programada. Eles demonstraram que plantas sem AtSerp1 foram mais comprometidas pela ação dos fungos *Botrytis cinerea* e *Sclerotinia sclerotiorum* nos mecanismos de morte celular programada do que em plantas que superexpressavam a proteína (LAMPL et al., 2013).

Ainda outro estudo envolvendo a proteína, AtSerp1, demonstrou que esta pode estar relacionada em vias de morte celular programada e estresse da planta por inibir, *in vitro*, uma metacaspase 9 (AtMC9) (VERCAMMEN et al., 2006), e, mais recentemente, estudos *in vivo*, mostraram potencial da proteína contra uma metacaspase 1 (AtMC9), sendo que ambas as enzimas têm sido relacionadas em vias de morte celular da planta (LEMA ASQUI et al., 2017). No arroz, uma outra serpina homóloga à AtSerp1 demonstrou regular negativamente a morte celular em processos de condições estressantes para a planta (BHATTACHARJEE et al. 2015).

Além de AtSerp1 outras duas serpinas também encontradas em *Arabidopsis*, a AtSRP2 e AtSRP3, demonstraram estar envolvidas em processos de regulação de crescimento da planta na presença de agentes alquilantes do DNA

(AHN et al. 2009). E, posteriormente, um estudo com outras duas serpinas de *Arabidopsis*, demonstrou que mutantes nocaute dos genes (*Atsrp4* e *Atsrp5*) que codificam para as serpinas AtSRP4 e AtSRP5 demonstraram aumento na morte celular induzida por resposta a radiação ultra violeta, quando em comparação com plantas do controle expressando essas proteínas normalmente. No estudo em questão, eles perceberam que a presença dessas serpinas reduziam a morte celular programada (BHATTACHARJEE et al. 2017). Esses dados em conjunto corroboram para o fato que essas serpinas, embora não tenham seus mecanismos totalmente elucidados, possam ter papel fundamental na regulação de enzimas e proteínas relacionadas a morte celular programada e infecção causado por patógenos.

4. MATERIAL E MÉTODOS

4.1. Análises *in silico* da sequência primária de TcSERPIN

A sequência completa dos resíduos de aminoácidos de TcSERPIN foi obtida a partir de busca no banco de dados do NCBI e confirmado no banco de dados do genoma do cacau, o CocoaGen DB (<http://cocoagendb.cirad.fr/>) (ARGOUT et al., 2007), segundo o número de acesso Tc08v2_p001150.1.

Predições referente à localização subcelular foram realizadas segundo TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) (EMANUELSSON et al, 2007) e PSORT (NAKAI & HORTON, 1999). Outras análises referentes a modificações pós traducionais foram realizadas com os programas SignalP 4.0 Server (<http://www.cbs.dtu.dk/services/SignalP-4.0/>) (PETERSEN et al, 2011) para verificação da presença de peptídeo sinal; NetPhos 2.0 (<http://www.cbs.dtu.dk/services/NetPhos-2.0/>) para identificar sítios putativos de fosforilação (Ser/Thr/Tyr) (BLOM et al, 1999); NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) (GUPTA et al., 2004) para supostos locais de N-glicosilação e NetAcet 1.0 Server (<http://www.cbs.dtu.dk/services/NetAcet-1.0/>) (KIEMER et al., 2004) para identificar sítios de acetilação.

As análises de massa molecular (mW) e ponto isoelétrico (pI) foram obtidas a partir da plataforma ExPASy (http://web.expasy.org/compute_pi/) (GASTEIGER et al, 2005).

A partir da sequência primária dos resíduos de aminoácidos de TcSERPIN foi realizado BLAST (ALTSCHUL et al, 1997) no National Center for Biotechnology Information (NCBI) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), para seleção de sequências homólogas de serpinas de plantas e outros organismos. Outras análises de alinhamento das estruturas tridimensionais foram realizadas também com o software COFATOR Server (<https://zhanglab.ccmb.med.umich.edu/COFACTOR/>) (ZHANG et al., 2017).

As sequências selecionadas foram submetidas a múltiplos alinhamentos por meio da plataforma Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) para análise de regiões de domínios e sítio ativo conservados. A partir dos alinhamentos foi gerado um dendrograma, visando análise filogenética das sequências primárias das proteínas selecionadas, segundo o programa MEGA 7, utilizando o método de *Neighbor-Joining*, com 1000 *BootStraps*. Para a construção do dendrograma, somente as sequências completas foram consideradas.

4.2. Modelagem por homologia de TcSERPIN e as cisteíno-proteases do cacau

A predição das estruturas tridimensionais (3D) de TcSERPIN e das cisteíno-proteases de *Theobroma cacao*, a RD19 (Tc02v2_p006310.1) e a RD21 (Tc02v2_p008580.1) selecionadas para o *docking*, foram realizadas com base na homologia da proteína por similaridade do modelo-alvo, utilizando o servidor Swiss-Model online (<https://swissmodel.expasy.org/interactive>) (ARNOLD et al., 2006).

A modelagem de TcSERPIN foi realizada utilizando como molde a estrutura cristalografada de AtSerp1 de *Arabidopsis thaliana* (Código PDB: 3LE2). As modelagens das duas cisteíno-proteases do cacau, a RD19 e RD21, cujas sequências primárias foram obtidas a partir do banco de dados do genoma do cacau (CocoaGenDB), foram realizadas segundo os moldes cristalografados de uma Cathepsina L de *Tenebrio molitor* (Código PDB: 3QJ3) e uma cisteíno-protease alergênica do pólen de *Ambrosia artemisiifolia* (Código PDB: 5EF4), respectivamente.

Todas as estruturas geradas ainda foram alinhadas utilizando os softwares TM-align (<https://zhanglab.ccmb.med.umich.edu/TM-align/>) (ZHANG & SKOLNICK, 2005) e PyMOL para avaliar as similaridades entre as estruturas 3D geradas e seus respectivos moldes. As estruturas tridimensionais das cisteíno-proteases foram ainda preparadas para a realização do docking, retirando o domínio pró-inibitório, segundo o software Discovery Studio 4.0.

A qualidade estereoquímica dos modelos obtidos foi avaliada utilizando os programas Procheck (LASKOWSKI et al., 1993) e Anolea (MELO & FEYTMANS, 1998) (<http://swissmodel.expasy.org/workspace/>). A visualização de todos os modelos gerados procedeu segundo os programas PyMol (The PyMOL Molecular Graphics System, Versão 1.5.0.4 Schrödinger, LLC.) e Discovery Studio 4.0.

4.3. Expressão e purificação de TcSERPIN recombinante

O clone recombinante foi obtido por meio da empresa GENONE, sendo que o gene sintético foi clonado em pET28a nos sítios de restrição *NcoI* e *XhoI*. A estratégia de clonagem do gene sintético utilizando a enzima *NcoI* possibilitou a exclusão da His-tag o que gerou uma proteína recombinante sem cauda de histidina.

A transformação bacteriana foi obtida a partir da estirpe de *Escherichia coli* Rosetta (*DE₃*) para absorção e incorporação do plasmídeo recombinante. Após a transformação, a estirpe bacteriana de *E. coli* Rosetta (*DE₃*) com o plasmídeo recombinante foi cultivada em meio LB (Luria-Bertani) e incubada a 37 °C sob agitação de 180 rpm até atingir a densidade ótica (OD) 600, aproximadamente, entre 0,7 e 1,0 η m de ABS. A indução da proteína recombinante foi obtida por adição de 0,4 μ mol.L⁻¹ de IPTG (Isopropil beta-D-tiogalactosídeo) durante 4 horas a 37 °C.

A cultura de células induzida foi submetida a centrifugação de 14.000 rpm e o precipitado bacteriano gerado foi ressuspensionado em tampão de lise contendo lisozima 50 μ g ml⁻¹ por 20 min. Após ressuspensão, o extrato total bacteriano foi submetido ao aparelho ultrassônico (pGEX 30) por 10 segundos contínuos e com intervalos de 25 segundos em repouso no gelo, com 70% de amplitude, para total ruptura das membranas e paredes celular. Após a lise, o material bacteriano foi centrifugado à rotação de 14.000 rpm, 4 °C por 20 minutos.

A proteína foi purificada a partir da fração insolúvel do extrato bacteriano, o sobrenadante foi descartado e o precipitado resultante foi submetido a diversas lavagens em tampão de lise contendo lisozima, nas mesmas concentrações citadas anteriormente, e, submetido ao ultrassonicador em cada lavagem. Após um total de 5 lavagens, o precipitado bacteriano foi ressuspenso em 20 mmol L⁻¹ de Tris HCL pH 7 e uréia 6 mol L⁻¹, contendo 5% de glicerol. Após a lise em ultrassonicador, o material bacteriano foi submetido à centrifugação a 14.000 rpm e o sobrenadante submetido à diálise gradual para remoção da uréia e re-enovelamento da proteína.

As análises para comprovação da expressão e purificação da rTcSERPIN foi realizado em gel SDS PAGE e corado com 0,01 % de Coomassie blue G-250.

4.4. Análise de Dicroísmo Circular (CD)

A análise de espectropolarímetro J-815 (JASCO) da proteína recombinante (TcSERPIN) em Dicroísmo Circular (CD), foi realizada na concentração de 0.15 µg µL⁻¹ da proteína em tampão Tris HCL (10 mmol L⁻¹, pH 7,0).

Para identificar a presença de conformações secundárias, a proteína foi condicionada em cubetas de quartzo de 1 mm e a varredura dos espectros da proteína foi obtida a partir da variação da absorvância entre 190-240 nm. Os dados foram coletados com velocidade de varredura de 50 nm min⁻¹ e em intervalos de 0.5 nm, com uma média de 6 varreduras a cada leitura, utilizando o software Spectra Measurement (Jasco). O percentual da estrutura secundária baseada no espectro de CD foi calculada a partir do software K2D3 (<http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3/>) (LOUIS-JEUNE et al., 2012).

A influência do pH associado a variação da temperatura também foi avaliada segundo *unfolding* (desnaturação por variação de 25-95 °C) e *refolding* (re-enovelamento da proteína por variação de 95-25° C). A análise de desenovelamento e re-enovelamento de rTcSERPIN foi caracterizada medindo as alterações da estrutura secundária da proteína em 217 nm em pH 7,0 e 214 nm em pH 5,4. Todas as análises de *unfolding* e *refolding* foram realizadas em triplicatas e a média dos sinais das triplicatas foram utilizadas para análise da estrutura secundária sob variação do pH e temperatura.

4.5. Ensaio da atividade inibitória de rTcSERPIN contra tripsina de *Sus scrofa* e papaína de *Carica papaya*

Para a análise das atividades inibitórias de rTcSERPIN contra a papaína de *Carica papaya* e a tripsina de *Sus scrofa* utilizou-se o substrato cromogênico BApNA ($1,2 \text{ mmol L}^{-1}$) (cloridrato de N-benzoil-D, L-arginina 4-nitroanilida), sendo a atividade da tripsina realizada segundo as adaptações de Sabotic *et al* (2012), em tampão Tris-HCL (50 mmol L^{-1}) e CaCl_2 (20 mmol L^{-1}), pH 7,4. Para a papaína a realização da atividade inibitória seguiu de acordo com as adaptações de Pirovani *et al* (2010), em tampão fosfato (50 mmol L^{-1} , pH 6,0) com adição de 2-mercaptoetanol (10 mmol L^{-1}) e EDTA (2 mmol L^{-1}).

Ambas as reações foram pré-incubadas por 10 minutos em 30°C contendo concentrações variadas do inibidor de 0,039; 0,047; 0,055; 0,063; 0,071 $\mu\text{g } \mu\text{L}^{-1}$ para a reação com a tripsina e 0,059; 0,071; 0,083; 0,095; 0,107 $\mu\text{g } \mu\text{L}^{-1}$ para a reação com a papaína. Sendo que, em cada reação, foram aplicadas $10 \mu\text{L}^{-1}$ das enzimas em concentrações fixas de $2,5 \mu\text{g } \mu\text{L}^{-1}$ para a tripsina e $50 \mu\text{g } \mu\text{L}^{-1}$ para a papaína crua do látex. Como controle para as duas reações foram utilizadas as respectivas enzimas juntamente com o substrato, sem a presença do inibidor. Após o tempo de incubação das enzimas com rTcSERPIN nos tratamentos contendo o inibidor, foi adicionado o substrato BApNA e realizada a leitura.

A reações foram monitoradas por 1 hora em temperatura de 30°C , com intervalos de 5 minutos, a 410 nm , em leitor de microplacas VersaMax (Tunable Molecular Devices, Silicon Valley, CA, EUA). A média das absorvâncias da atividade inibitória de rTcSERPIN foi registrada a partir de cálculos em Excel, sendo que a porcentagem de 100 % da atividade residual da tripsina e da papaína foram determinadas a partir das leituras do controle (sem inibidor).

4.6. Análise de biologia de sistemas

Para ampliar o conhecimento da relação de TcSERPIN com possíveis alvos endógenos foi gerada uma rede de interação utilizando como base o banco de dados de AtSerp1 de *Arabidopsis thaliana*. Os dados para confecção da rede foram gerados segundo o banco de dados STRING 9.05 (<http://string-db.org>). E os

resultados obtidos foram analisados segundo o software Cytoscape versão 2.8.3 [43].

Os padrões para a construção levou em consideração mais de 50 interações, nível de significância de 0,6 e adição de nós até que uma rede saturada fosse obtida. O *plugin* MCODE foi usado para agrupar as proteínas que tiveram maiores relações entre si, a partir do maior percentual de interação entre elas. O *plugin* BiNGO foi utilizado para analisar os processos biológicos, os quais foram observados na rede em geral. Por fim, somente o cluster mais significativo foi escolhido para a análise.

4.7. Docking molecular

4.7.1. Docking entre TcSERPIN e as serino- e cisteíno-proteases

Para a realização do acoplamento (*docking*) entre os modelos 3D de TcSERPIN (receptor) e as moléculas-alvo (ligantes) foram selecionadas três estruturas cristalografadas, duas serino-proteases do pâncreas de *Bos taurus*, a α -quimotripsina e a α -tripsina (Código PDB: 1YPH e 1S0Q, respectivamente) e uma cisteíno-protease, a papaína de *Carica papaya* (Código PDB: 1CVZ). Também foram selecionadas duas cisteíno-proteases de *T. cacao*, a RD19 e RD21 modeladas por homologia com a papaína de *Tenebrio molitor* (Código PDB: 3QJ3) e *Ambrosia artemisiifolia* (Código PDB: 5EF4), respectivamente. O acoplamento do inibidor e das proteases-alvo foram obtidas por meio do software ClusPro 2.0 (<http://cluspro.bu.edu/>), um programa de ancoragem de proteínas (KOZAKOV et al., 2017).

As análises dos acoplamentos levou em consideração estimativa de níveis de energia (van der Waals, forças eletrostáticas e hidrofóbicas), o tamanho de cada cluster gerado pelo programa e interações presentes no sítio ativo das proteínas envolvidas. Finalmente, o melhor modelo do complexo Receptor-Ligante foi selecionado e analisado utilizando os programas PyMol e Discovery Studio 4.0.

4.7.2. *Docking* entre TcSERPIN e a heparina de sódio

O *docking* entre TcSERPIN (receptor) e a heparina (ligante) foi realizado para avaliar a possibilidade de ocorrência de ligação entre a serpina do cacau e o pentassacarídeo. Para isso, foi usado como base regiões de afinidade de proteínas homólogas da superfamília das serpinas que possuem afinidade com esse ligante, como o Cofator de Heparina II e a Antitrombina.

A heparina foi obtida no PubChem database (<http://pubchem.ncbi.nlm.nih.gov>). A estrutura em 3D desse composto foi obtida por meio do programa Marvin, copiada em Similies e salvo em formato Mol 2. O receptor e o ligante foram preparados por meio do software AutoDockTools 1.5.6 (SANNER et al., 1996) e o ajuste da caixa (grid Box) ao sítio ativo da proteína foi configurada manualmente seguindo as recomendações do manual do software (SANNER et al., 1996; TROTT & OLSON, 1996). As moléculas foram salvas em formato pdbqt e o AutoDock Vina foi utilizado para executar *Docking Scoring* do complexo formado pelo receptor–ligante (TROTT & OLSON, 2010). Os cálculos de *docking* foram analisados visando obtenção dos valores de energia (Kcal/Mol) de afinidade para cada conformação receptor-ligante. O programa PyMol foi utilizado para avaliação do complexo mais adequado segundo as análises de energia de afinidade e sítio ativo da proteína receptora. A conformação com melhor encaixe e valor de energia livre foram selecionadas. As imagens finais dos complexos e o mapa 2D foi gerado utilizando o programa Discovery Studio 4.0.

4.8. Extração de proteínas nativas totais das folhas de *T. cacao*

A extração de proteínas nativas totais das folhas de *Theobroma cacao* da variedade Scavina 6 foi realizada utilizando 1g de material vegetal macerado e ressuspenso em 5 mL de Tampão de Extração (Tris-HCl 10 mmol L⁻¹ pH 7,5, Triton X-100 1 %). As amostras foram agitadas em vórtex por 10 min, sonicadas (8 pulsos de 3 s, com intervalos de 10 s, amplitude 70%) e centrifugadas (5000 rpm por 20 min, 4 °C).

O precipitado resultante da etapa anterior foi submetido a uma segunda extração. Ao sobrenadante obtido foram adicionados 0,4 volumes de *tert*-Butanol; 1/10 de NaAc, 3 mol L⁻¹, pH 4,5. Posteriormente, as amostras foram mantidas em

gelo por 30 minutos, sendo agitadas a cada 10 minutos em vórtex e centrifugadas a 5000 rpm por 10 min, 4 °C. As amostras contendo proteínas totais foram estocadas em 4 °C para posterior captura por sepharose ativada com CNBr acoplada com a serpina recombinante.

4.9. Armadilha da protease por CNBr-Sepharose

Para montagem do sistema de captura da protease-alvo foi necessário realizar o acoplamento de TcSERPIN recombinante a SepharoseTM 4 Fast Flow ativada por CNBr (GE Healthcare) seguindo as instruções do fabricante. A partir do acoplamento do inibidor foi realizada a captura da protease-alvo. O extrato de folhas de *T. cacao* foi incubado com a resina contendo o inibidor acoplado por 40 minutos em temperatura ambiente, sob leve agitação, na presença de tampão de atividade (tampão fosfato de potássio, 50 mmol L⁻¹, pH 6,0). Após a captura das proteases presentes no extrato da folha, a resina foi lavada 3 vezes com tampão de atividade contendo NaCl 0,5 mmol L⁻¹. As proteases que, possivelmente, interagiram com o inibidor (TcSERPIN) foram eluídas com 50 µL⁻¹ de glicina, 50 mmol L⁻¹, pH 2,9, por 5 minutos, alternando para volume igual de Tris-base de 10 mmol L⁻¹, para equilíbrio, até um volume total de 200 µL⁻¹. Entre cada etapa, a resina foi recuperada por centrifugação (3000 rpm, por 30s, temperatura 4 °C). Os passos mediante a captura das proteínas foram realizados segundo adaptações de Pirovani *et al* (2010).

As proteases capturadas foram analisadas para avaliação da atividade da protease através de um ensaio qualitativo. O ensaio da avaliação da atividade foi possível através de análise em gel de SDS-PAGE. Após a migração das proteínas, os géis foram lavados em Triton X-100, à temperatura ambiente, por 30 min para remoção do SDS. Posteriormente, o gel foi recoberto com poliacrilamida a 7,5 % de gel de gelatina a 0,1% na presença de tampão de atividade da protease em temperatura de 37 °C por 48 h (MICHAUD *et al.*, 1996; PIROVANI *et al.*, 2002, 2010). Após o tempo de incubação, os géis foram então corados com 0,01 % de Coomassie blue G-250.

5. RESULTADOS

5.1. Análise *in silico* da sequência primária de TcSERPIN

O gene que codifica para a sequência correspondente da proteína TcSERPIN (Tc08v2_p001150.1) de *Theobroma cacao* encontrado no CocoaGen DB está localizado no cromossomo 8, e codifica uma proteína com massa estimada de 42,5 kDa e pI putativo de 5.88. As análises das modificações pós-traducionais revelaram 21 sítios possíveis de fosforilação e dois sítios de glicosilação (Figura 4). As análises *in silico* não revelaram presença de peptídeo sinal, porém, dados de predição e alinhamento com sequências homólogas realizadas segundo o COFATOR Server demonstraram que outras proteínas da família possam ser encontradas no apoplasto. Outros dados da literatura também reforçam essa possibilidade (LAMPL et al., 2010; VERCAMMEN et al., 2006).

5.2. Alinhamento de TcSERPIN e sequências homólogas

O alinhamento entre TcSERPIN e outras serpinas homólogas obtidas a partir do banco de dados do NCBI por meio de BLSTp demonstrou conservação dos resíduos da região do RCL, onde é possível visualizar a presença típica de uma glicina na posição P15, comum a todas as serpinas, seguidas pela presença de treonina/serina em P14, glutamato em P13 seguidos de alta conservação de alaninas de P9 a P12, presentes em quase todas as serpinas inibitórias, sublinhado no alinhamento (Figura 4). Esse padrão é observado na serpina inibitória de humanos (Serpina B6) bem como nas três serpinas inibitórias de plantas (TcSERPIN de *Theobroma cacao*, BSZx de *Hordeum vulgare* e AtSerpina1 de *Arabidopsis thaliana*). A exceção desse padrão, no entanto, é da única serpina não-inibitória de humanos, Maspin.

Maspin possui uma glicina na posição 14 e não possui uma sequência contínua de alaninas a partir de P12 (Figura 4), situação comum em serpinas que não possuem a capacidade inibitória (HOPKINGS et al., 1993; IRVING et al., 2000). As três serpinas de plantas, como observado no alinhamento, possuem alta

conservação de resíduos do RCL e apresentam motivo conservado existente somente em plantas (Figura 4).



Figura 4: Alinhamento de TcSERPIN e outras sequências homólogas: Em destaque a região do RCL das cinco serpinas: TcSERPIN de *Theobroma cacao*, AtSerp1 de *Arabidopsis thaliana*, BSZx de *Hordeum vulgare*, SerpinB6 e Maspin de *Homo sapiens*; em amarelo e verde, os sítios de fosforilação e glicosilação de TcSERPIN, respectivamente; em azul, regiões de motivo conservado nas cinco serpinas; em rosa, motivo conservado somente em serpinas de plantas; em cinza, regiões conservadas somente nas serpinas inibitórias. Sublinhado, o RCL.

Ainda para avaliar a relação de homologia e identidade das sequências primárias de serpinas de plantas e outros organismos, foi gerado um dendrograma, onde é possível visualizar uma separação distinta entre as serpinas de plantas com outros organismos, como mamíferos, insetos, alguns microrganismos e certos vírus (Figura 5).

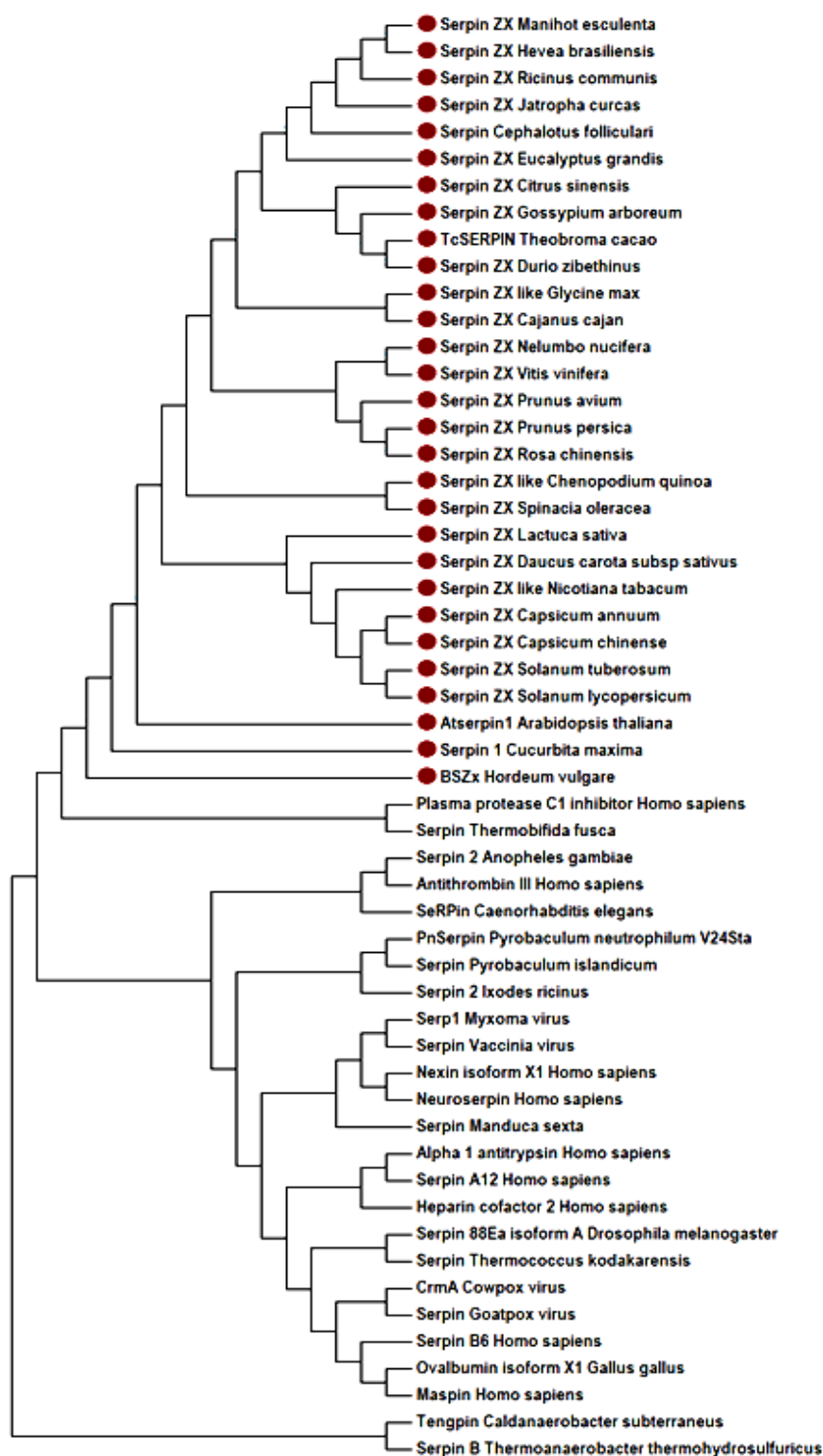


Figura 5. Dendrograma do alinhamento de TcSERPIN com sequências homólogas: As marcações em vermelho mostram exclusivamente serpinas de plantas, incluindo TcSERPIN.

Observando atentamente a região do RCL das serpinas de plantas selecionadas no dendrograma verifica-se o padrão de conservação dessas serpinas dentro do clã, com a presença de uma glicina na região P15 (*hinge region*) e uma

treonina em P14 em todas as serpinas do alinhamento, seguidas de glutamato ou lisina (P13) e uma sequência significativa de alaninas de P9 - P12 em quase todas as serpinas mostradas no alinhamento. A análise de outros resíduos de aminoácidos próximos à região do RCL até a região C-Terminal mostra uma região de conservação dos resíduos apresentados e de extensa homologia (Figura 6).

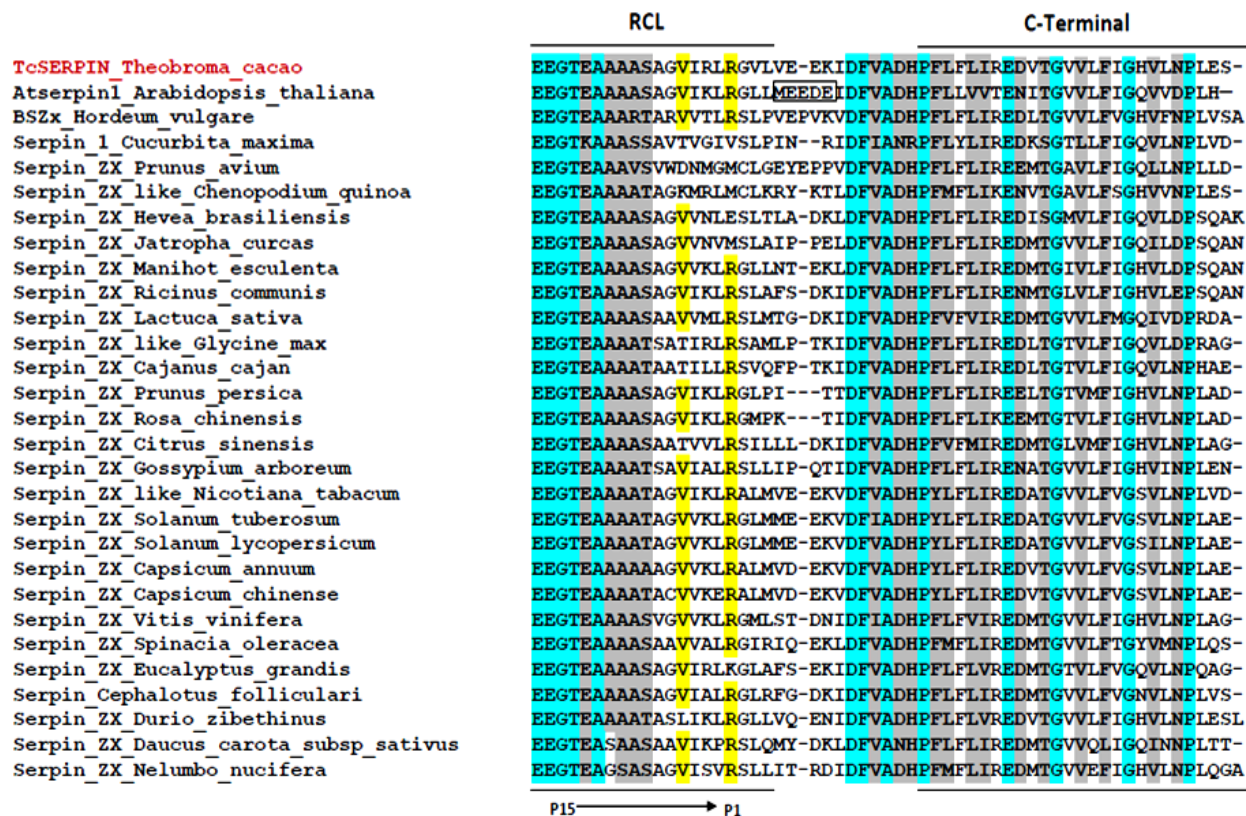


Figura 6. Alinhamento do RCL de serpinas de plantas mostradas no dendrograma: As marcações em azul mostram resíduos de aminoácidos conservados em todas as serpinas; em cinza, regiões de extensa homologia entre as serpinas, da região do RCL em direção a região C-Terminal; em retângulo, região de motivo exclusiva de *Arabidopsis thaliana*; em amarelo, possíveis regiões de clivagem do RCL. Em destaque TcSERPIN (vermelho) e o RCL de todas as serpinas do alinhamento de P15 - P4'.

5.3. Modelagem

A modelagem de TcSERPIN foi realizada segundo o software Swiss Model online, utilizando como molde a estrutura cristalografada em raio-X de AtSerp1 de *Arabidopsis thaliana* (3LE2 PDB) em seu estado nativo estressado, o que viabilizou a obtenção do modelo 3D com 67,78 % de identidade e 99 % de cobertura (Figura 7a).

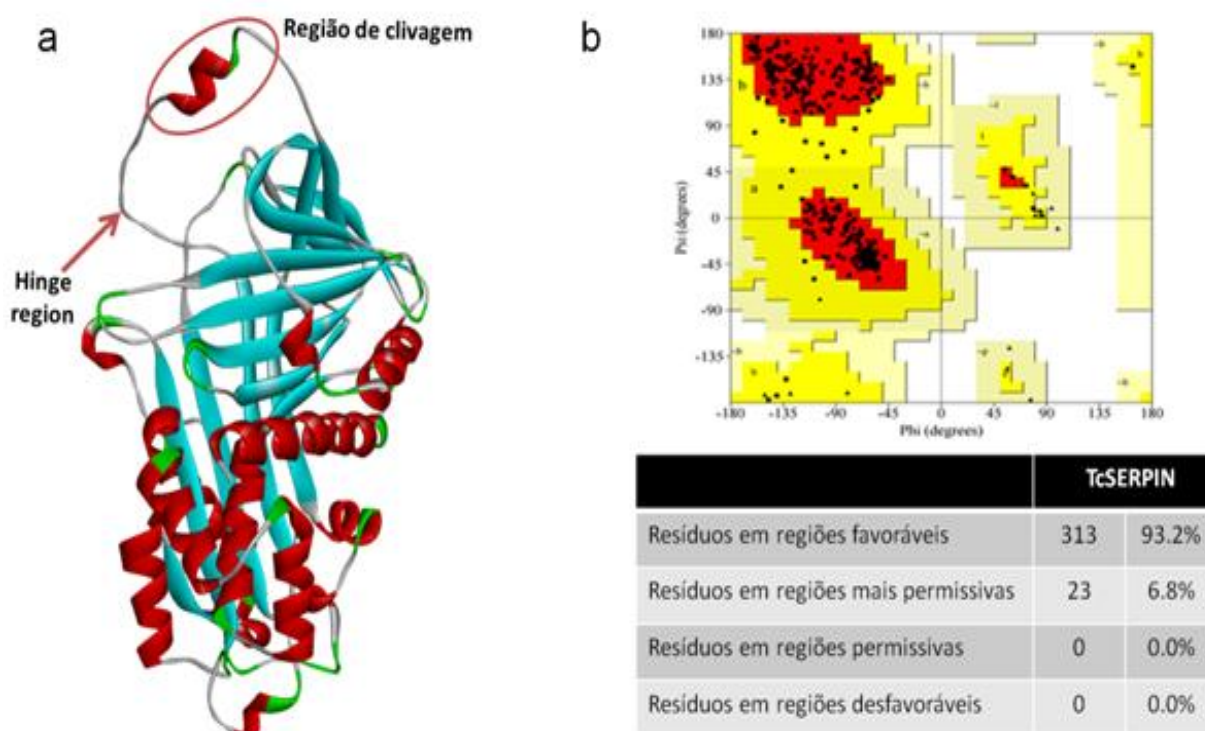


Figura 7: Estrutura tridimensional de TcSERPIN a partir do molde AtSerp1 de *Arabidopsis thaliana* (Código PDB: 3LE2): Em **a**, TcSERPIN, mostrando em destaque regiões conservadas importantes para o mecanismo de inibição; Em **b**, gráfico de Ramachandram validando o modelo tridimensional.

A modelagem das estruturas das proteases-alvo de *Theobroma cacao* tendo como molde uma Cathepsina L de *Tenebrio molitor* (Código PDB: 3QJ3) obteve 38 % de identidade e 80 % de cobertura para RD19 (Figura 8a) e uma cisteíno-protease alergênica do pólen de *Ambrosia artemisiifolia* (Código PDB: 5EF4) com 45,5 % de identidade e 69 % de cobertura para RD21 (Figura 8b).

A validação dos modelos gerados de TcSERPIN (Figura 7b) e das cisteíno-proteases (Figura 8c, d, e) de *T. cacao* usando o software do Procheck, mostrou um diagrama de Ramachandram com 100 % dos resíduos de aminoácidos em regiões energeticamente favoráveis, indicando que os modelos 3D são energeticamente adequados.

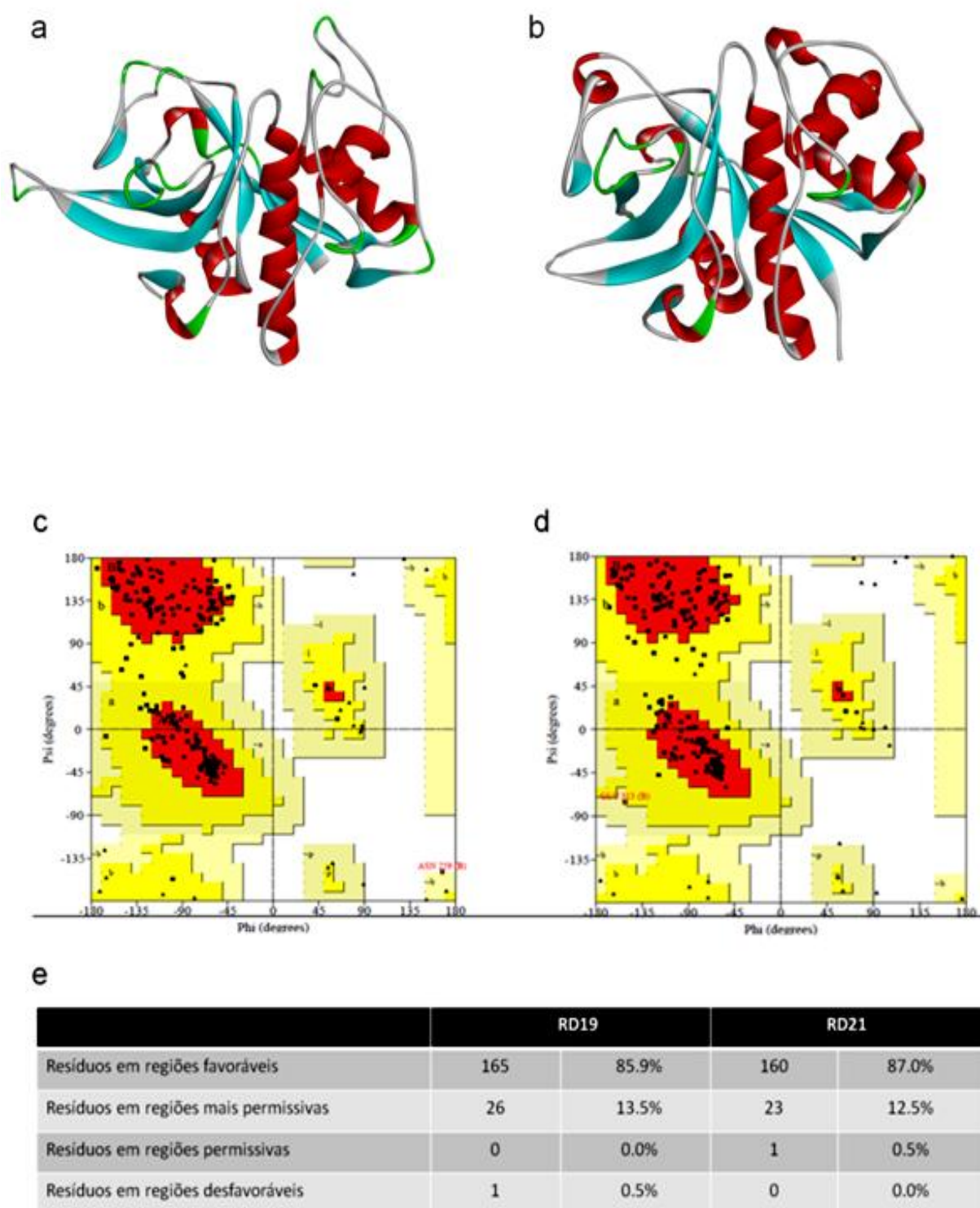


Figura 8. Estrutura tridimensional das cisteíno proteases do cacau: Em **a**, RD19 segundo molde de uma Cathepsina L de *Tenebrio molitor* (Código PDB: 3QJ3); em **b**, RD21 segundo molde de cisteíno-protease alergênica do pólen de *Ambrosia artemisiifolia* (Código PDB: 5EF4); em **c** e **d**, gráfico de Ramachandram de RD19 e RD21, respectivamente, mostrando que os modelos são energeticamente favoráveis; em **e**, valores dos resíduos de aminoácidos dos gráficos apresentados em cada região.

5.4. Expressão e purificação de TcSERPIN recombinante

A análise de TcSERPIN recombinante expressa em *Escherichia coli*, Rosetta DE3, por tempo de indução de 4 horas, em 37° C e 0,4 $\mu\text{mol L}^{-1}$ de IPTG, revelou uma banda de aproximadamente 42,5 kDa em gel SDS-PAGE. A banda correspondente à TcSERPIN acumula em maior proporção na fração insolúvel, embora uma pequena quantidade seja mostrada na fração solúvel.

A proteína recombinante sem cauda de histidina foi purificada a partir da fração insolúvel, seguindo uma estratégia desenvolvida a partir de lavagens com tampão de lise, sonicação e centrifugações contínuas para eliminação da maior parte dos contaminantes, o que possibilitou grande rendimento da proteína rTcSERPIN (Figura 9).

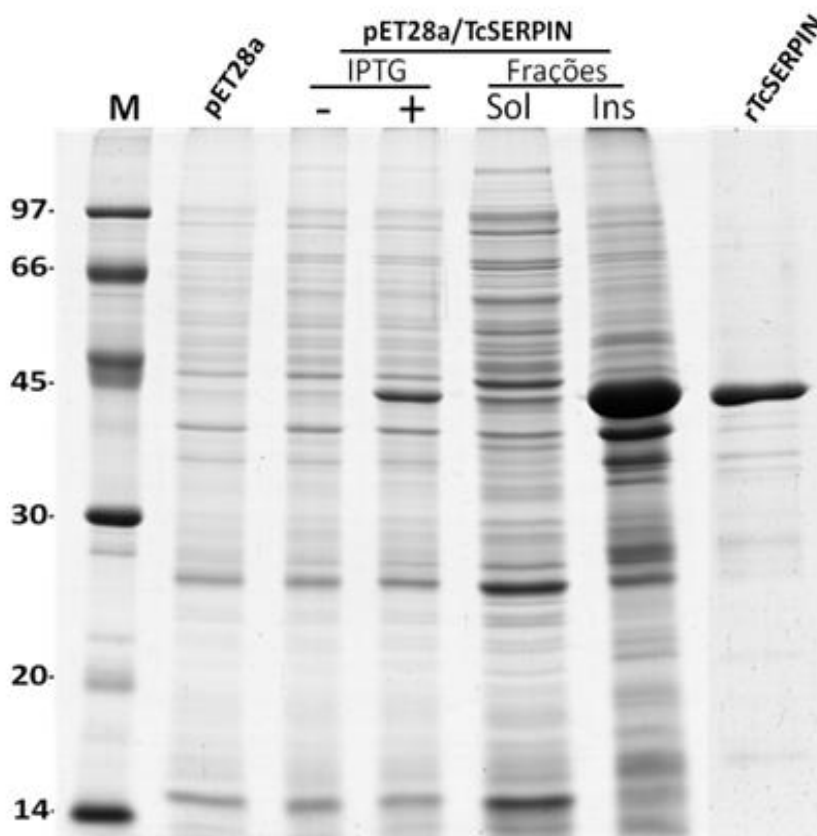


Figura 9: Expressão e purificação de TcSERPIN recombinante: Linha 1 (M), marcador; linha 2, pET28a (controle negativo); linha 3, pET28a contendo gene *TcSERPIN* sem indução por IPTG (controle negativo); linha 4, 5 e 6, análise da expressão de TcSERPIN representando o extrato total, fração solúvel e fração insolúvel do extrato bacteriano após adição por IPTG com 4 hs de indução (controle positivo); linha 7, TcSERPIN recombinante purificada a partir da fração insolúvel.

5.5. Análise da estrutura secundária de TcSERPIN em Dicroísmo Circular (CD)

As análises dos espectros de TcSERPIN em pH 7.0 no CD mostraram que a proteína possui uma variação entre folhas alphas e betas, sendo que o menor pico (-19.7 mdeg) ocorreu em torno de 216 nm a 25 °C. A análise do percentual da estrutura secundária segundo o software K2D3 demonstrou que a proteína recombinante possui 45 % de α -hélice e 14 % de folhas- β (Figura 10).

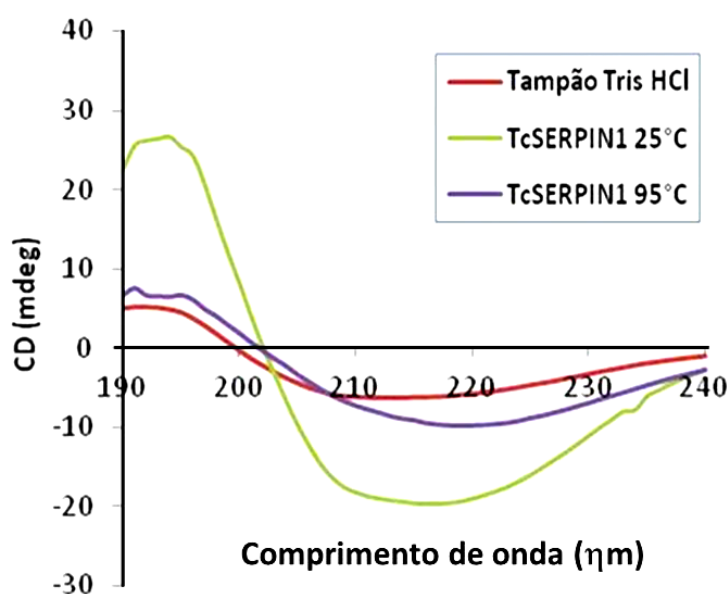


Figura 10. Avaliação dos espectros da estrutura secundária de TcSERPIN: Linha verde e azul do gráfico, mostram a proteína recombinante em tampão Tris HCL 10 mmol L⁻¹ em 25 e 95 °C, respectivamente; linha vermelha, somente o tampão Tris HCL 10 mmol L⁻¹ em 25 °C (controle).

As análises mediante a variação de temperatura e pH da proteína recombinante, tanto do desenovelamento (*Unfolding*) como no re-enovelamento (*Refolding*) demonstrou que TcSERPIN não suporta grandes variações de temperatura e não re-enovela quando submetida ao resfriamento, o que significa que a proteína não retorna a sua conformação inicial. No entanto, a proteína demonstrou variação do TM quando submetida a variação de temperatura em pH 5.4 e pH 7.0. A variação de temperatura em pH 5.4 leva a uma perda de sinal da proteína após 37° C, quando que, em pH 7.0 a perda de sinal ocorre após 66° C (Figura 11).

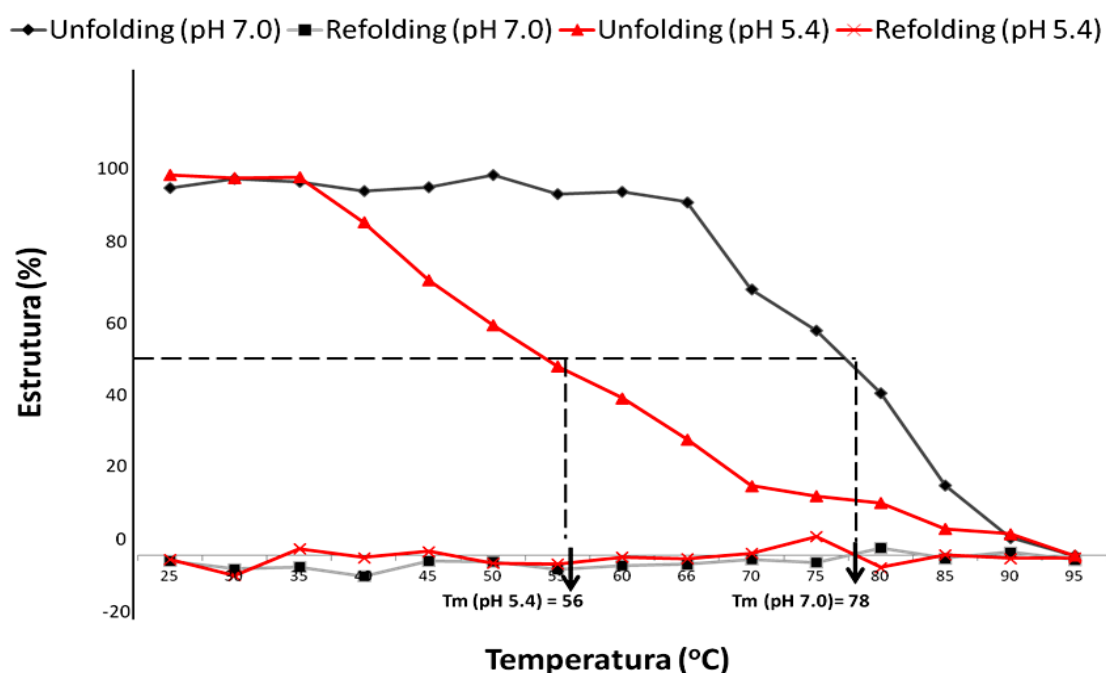


Figura 11. Unfolding e Refolding de TcSERPIN em pH 7.0 e 5.4: Linha preta, pH 7.0; linha vermelha, pH 5.4. No eixo X a variação da temperatura; No eixo Y, o percentual de estrutura da proteína recombinante.

5.6. Análise da atividade inibitória de rTcSERPIN contra tripsina suína e a papaína de *Carica papaya* *in vitro*

A rTcSERPIN demonstrou atividade inibitória frente a tripsina suína (*Sus scrofa*) e a papaína do mamão (*Carica papaya*). Nas análises que se seguiram, tanto contra tripsina como a papaína, rTcSERPIN foi incubada individualmente em concentrações variadas com cada protease por 10 minutos antes da adição do substrato cromogênico (BAPNA).

A proteína recombinante mostrou atividade inibitória de até 54,6 % contra a papaína (Figura 12a) e 27,85% contra a tripsina (Figura 12b). A medida que houve aumento da concentração do inibidor, ocorreu redução da atividade das proteases, sendo que, o maior percentual inibitório do inibidor foi alcançado nas concentrações de 0,1 e 0,07 $\mu\text{g } \mu\text{L}^{-1}$, respectivamente.

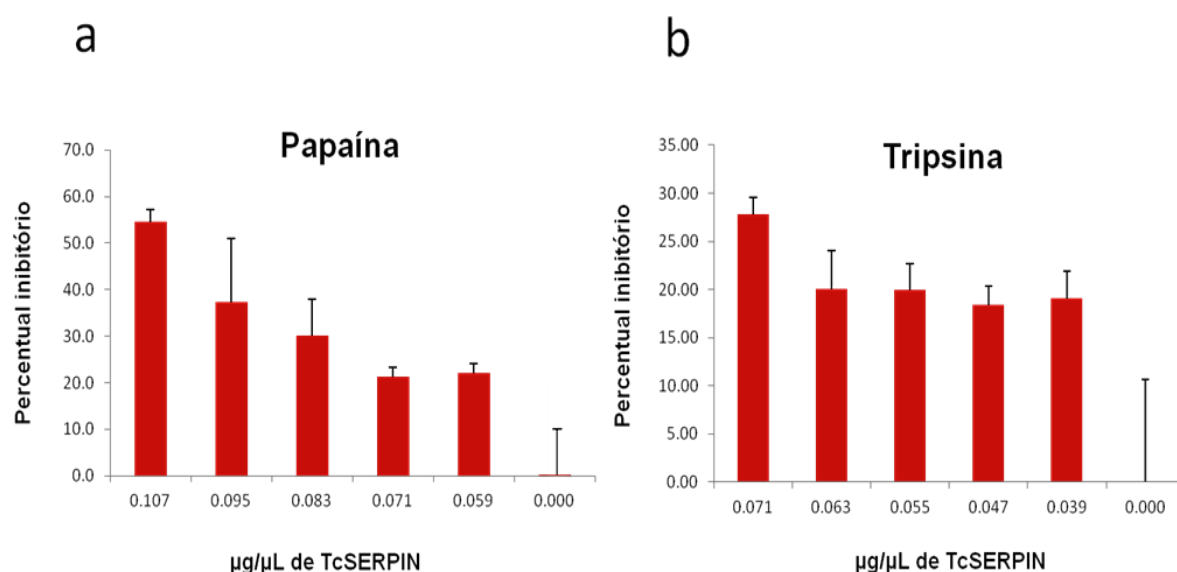


Figura 12: Teste de atividade de rTcSERPIN contra papaína do mamão e tripsina suína: Em **a** e **b**, avaliação do percentual de inibição de rTcSERPIN contra papaína e a tripsina; o eixo X mostra as concentrações do inibidor em $\mu\text{g } \mu\text{L}^{-1}$; o eixo Y, percentual de inibição após 30 minutos de interação entre o inibidor e a protease.

5.7. Análise *in silico* da interação de TcSERPIN com proteases do tipo serino e cisteíno

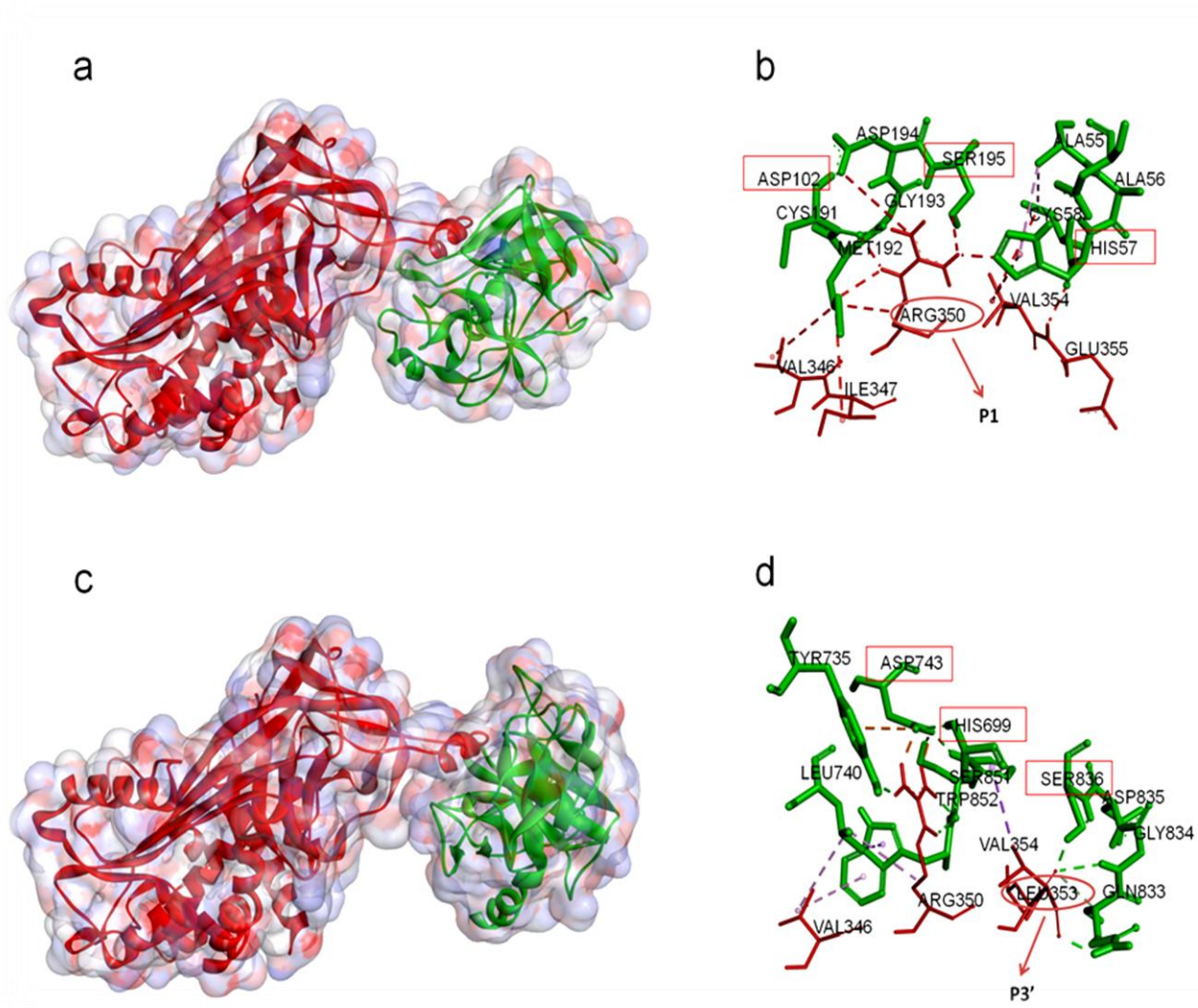
O acoplamento entre TcSERPIN e as estruturas cristalografadas da α -quimotripsina e α -tripsina bovinas (1S0Q PDB) e a papaína do mamão (1CVZ PDB) resultou em 10 conformações possíveis, sendo, no entanto, escolhidas aquelas mais favoráveis, segundo estimativas dos níveis de energia e o tamanho de cada cluster, bem como os locais de sítio ativos das proteínas envolvidas.

Todas as interações ocorreram em locais esperados de sítio ativo das proteases envolvidas e de TcSERPIN, mas com uma pequena variação do sítio de clivagem do inibidor (Figura 13).

O modelo de acoplamento de TcSERPIN com a quimotripsina (Figura 13 a, b) e tripsina (Figura 13 c, d) mostra interações no sítio ativo tanto da protease como da serpina. Durante a interação com a quimotripsina é possível visualizar uma serina 195 interagindo com uma arginina 350 (região de P1 segundo o alinhamento e possível região de clivagem do RCL) da serpina. No entanto, a interação da tripsina com a serpina mostra a serina 195 interagindo com uma leucina 353 (região P3' de TcSERPIN). Outras interações como a presença da tríade catalítica tanto da

quimotripsina (SER 195, HIS 57 e ASP 102) como da tripsina (SER 836, HIS 699 e ASP 743) também são visíveis durante a interação (Figura 13b, d).

A interação de TcSERPIN com a papaína do mamão, no entanto, ocorreu entre uma valina na posição 346 (região P5 de TcSERPIN) e a cisteína 25 da papaína (Figura 13e, f), onde, ao invés, da presença de uma arginina 350, é possível observar uma interação entre a valina 346 da serpina e a cisteína 25 da protease, o que pode significar uma variação do sítio de clivagem do inibidor. A arginina 350, porém, também pode ser vista na imagem interagindo com uma serina 21 e uma cisteína 22 da papaína. Também é possível visualizar a presença da tríade catalítica da papaína (CYS 25, HIS 159 e ASN 175) (Figura 13f).



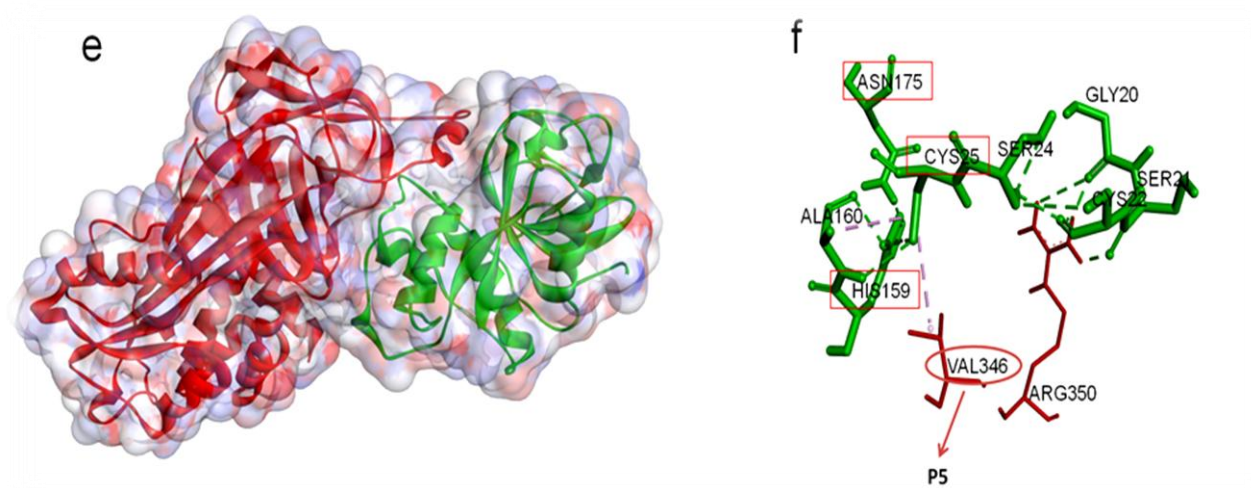


Figura 13. Docking entre TcSERPIN e as serino- e cisteíno-proteases: Em **a**, **c**, o acoplamento entre TcSERPIN com a α -quimotripsina e a α -tripsina (*Bos taurus*), respectivamente; em **e**, o acoplamento entre TcSERPIN e a papaína (*Carica papaya*); **b**, **d**, **f**, mostram o destaque das regiões de sítio ativo da serpina com as proteases (quimotripsina, tripsina e papaína). As moléculas em verde representam o ligante (as proteases) e em vermelho, o receptor (a serpina).

5.8. Rede de interação da TcSERPIN

A análise da rede de interação retornou três clusters. Uma análise detalhada relacionada com cada cluster objetivou selecionar somente o cluster 1, com melhor pontuação e valor estatístico que estava diretamente relacionada à serpina e as proteases.

O cluster 1 retornou com proteases envolvidas em processos proteolíticos e de estresse, relacionados a dessecação. Com base nesse cluster, e de sugestões da literatura (LAMPL et al., 2010) duas proteases candidatas, a RD19 e a RD21, relacionadas a processos de dessecação e morte celular programada foram selecionadas para a análise de *docking* de TcSERPIN e das cisteíno-proteases do cacau (Figura 14).

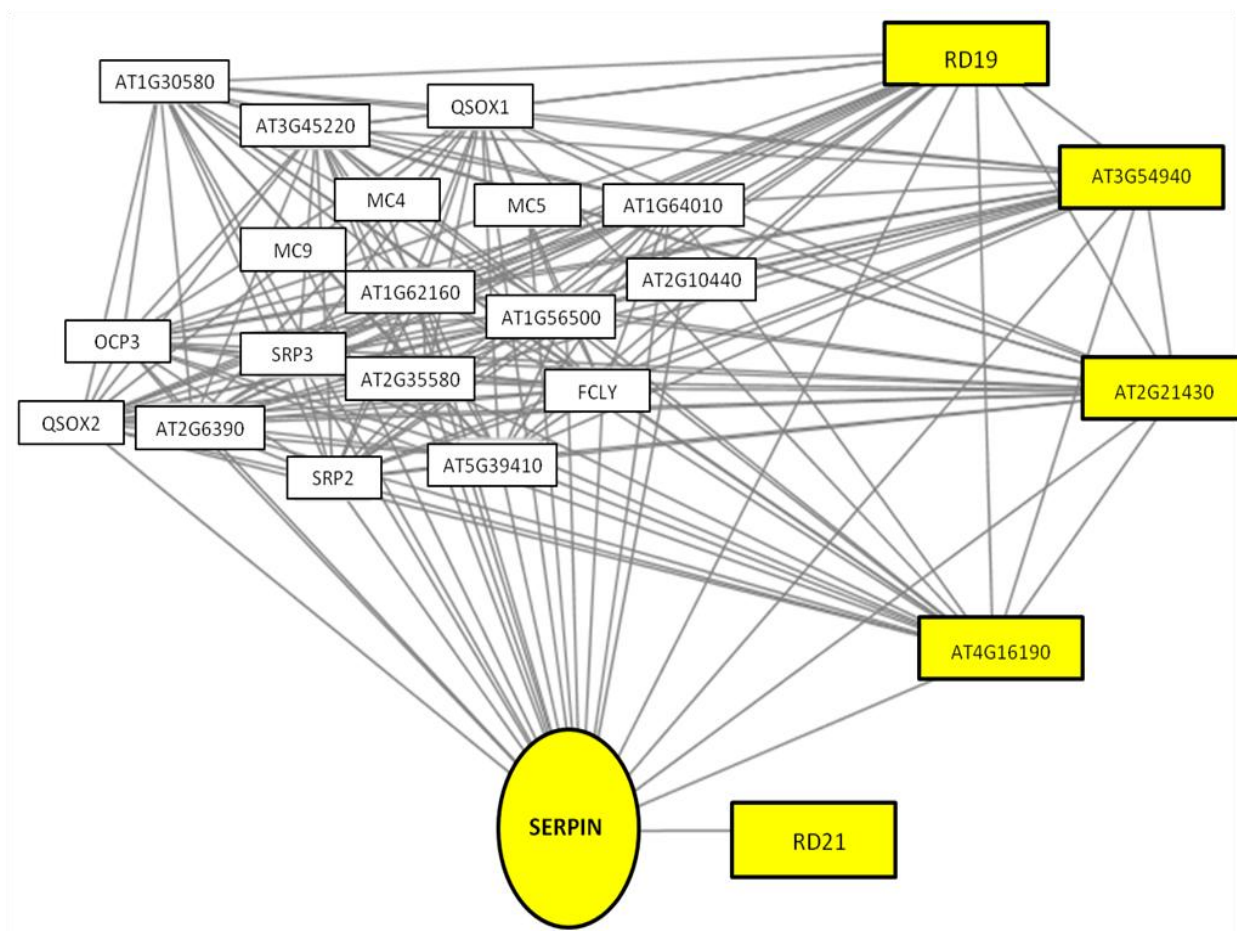


Figura 14. Cluster 1 mostrando as principais proteínas das vias relacionadas a TcSERPIN: As cores em amarelo mostram as proteínas com os melhores valores estatísticos e maior relação.

5.9. Análise *in silico* da interação de TcSERPIN com cisteíno-proteases de *Theobroma cacao*

Com base em informações da literatura (LAMPL et al., 2010) e a partir de análises de rede de interação de TcSERPIN baseados em sua homóloga, AtSerpin1 de *Arabidopsis thaliana* foi selecionado duas papaínas de *Theobroma cacao*, a RD19 (Cod.:Tc02v2_p006310.1) e a RD21 (Cod.:Tc02v2_p008580.1) no banco de dados online do CocoaGenDB.

Após as modelagens, as papaínas foram preparadas retirando o sítio pró-inibitório, um trecho longo de resíduos de aminoácidos que recobre o sítio ativo da protease, e, ao qual, sem a devida remoção impede a ação da papaína (BETON et al., 2012).

Os sítios de interação de TcSERPIN e as duas papaínas de *T. cacao* variaram entre a alanina 344 (região P7) e valina 346 (região P5) da serpina e as cisteínas do sítio ativo das peptidases, RD19 (Figura 15a, b) e RD21 (Figura 15c, d), respectivamente. Como pode ser visto na imagem, o acoplamento entre TcSERPIN e a RD21 apresentou resultado semelhante ao da papaína cristalografada de *Carica papaya*, em que as interações do tipo ligações de hidrogênio ocorreram entre a região P5 da serpina e a cisteína 25 do sítio catalítico da papaína (Figura 13e, f). As interações pertinentes ao sítio catalítico podem ser conferidas na imagem (Figura 15b, d).

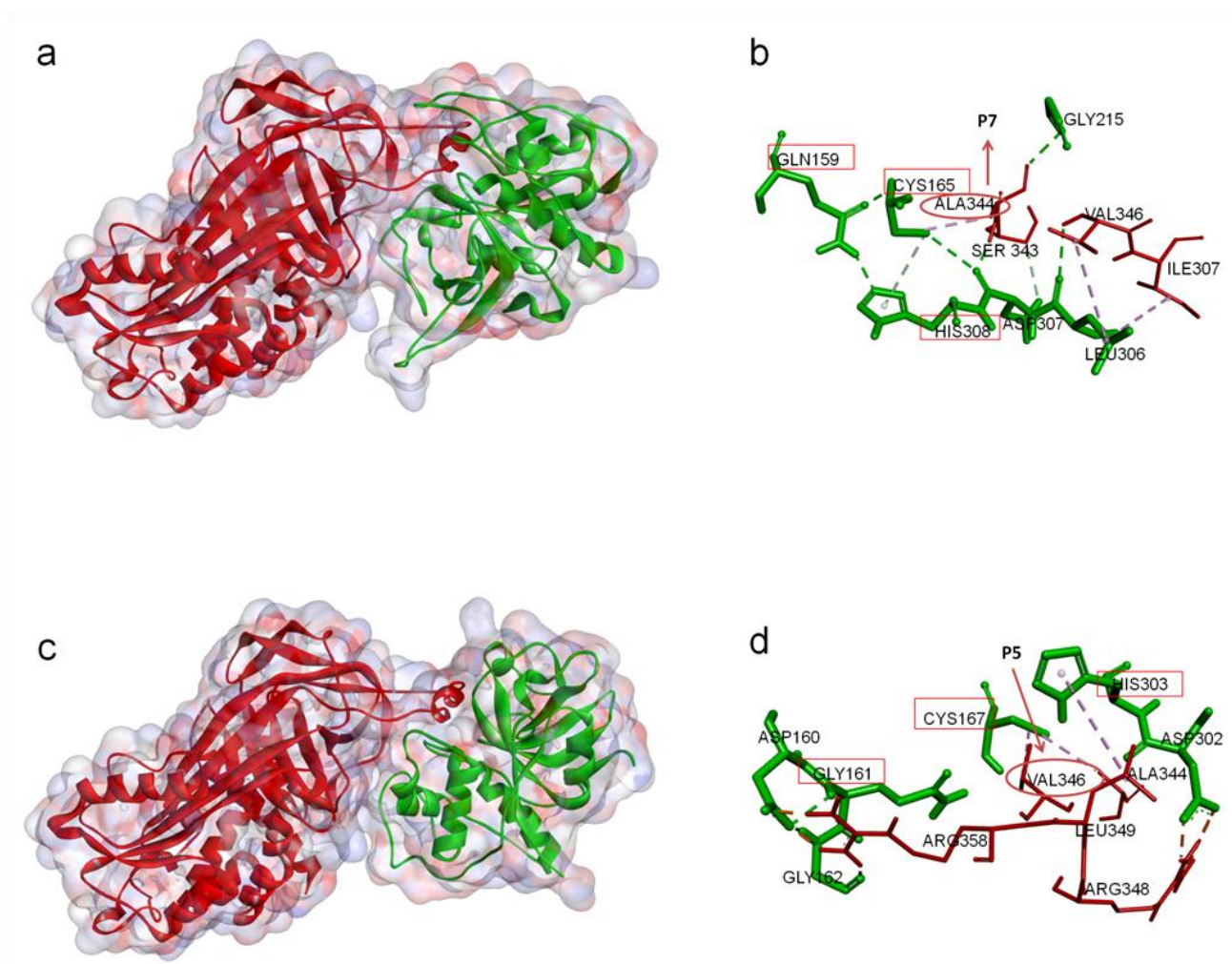
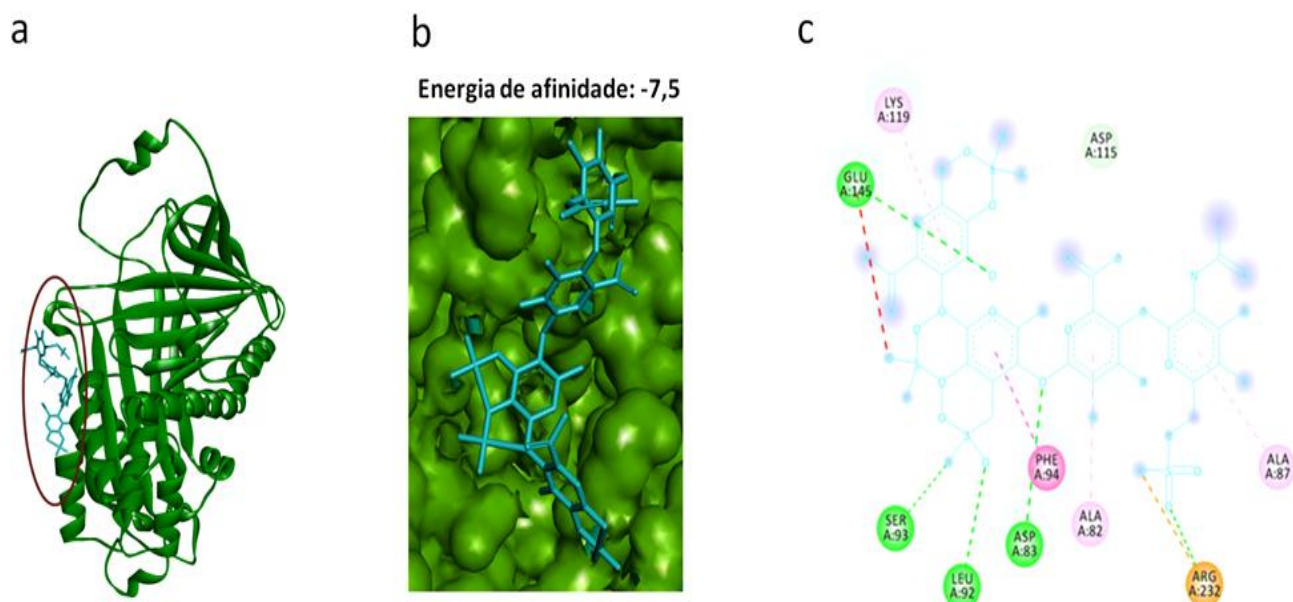


Figura 15. Docking entre a TcSERPIN e as papaínas RD19 e RD21 de *T. cacao*: Em **a**, o acoplamento entre TcSERPIN e RD19; em **c**, o acoplamento de TcSERPIN e RD21; em **b** e **d**, o destaque das regiões de sítio ativo da serpina e das proteases (RD19 e RD21, respectivamente). As moléculas em verde representam o ligante (as proteases) e em vermelho, o receptor (a serpina).

5.10. Análise da interação da heparina de sódio com TcSERPIN em sítios similares da Antitrombina-III

As análises de *docking* de TcSERPIN (receptor) e a heparina (ligante) levou em consideração o sítio de ligação na hélice D, característico de inibidores como a Antitrombina e o Cofator de heparina II.

A heparina de sódio (Código: 22833565) de peso molecular de $1039.8 \text{ g mol}^{-1}$ foi selecionada por meio do banco de dados do PubChem. A avaliação da interação entre a serpina e a heparina demonstrou que esse ligante possui afinidade com a serpina do cacau em várias regiões. No entanto, foi selecionado somente as regiões que possuíam menor energia de afinidade (-7,0 - 6,0) e sítios de ligação à heparina semelhante ao da Antitrombina. Devido ao fato de que a heparina obtida possuía baixo peso molecular, foi analisado três regiões de ligação entre a hélice D e o RCL, como proposto para serpinas com afinidade à heparina. As três regiões de ligação foram favoráveis em níveis de energia de ligação (energia mais negativa) e sítio ativo da proteína (Figura 16).



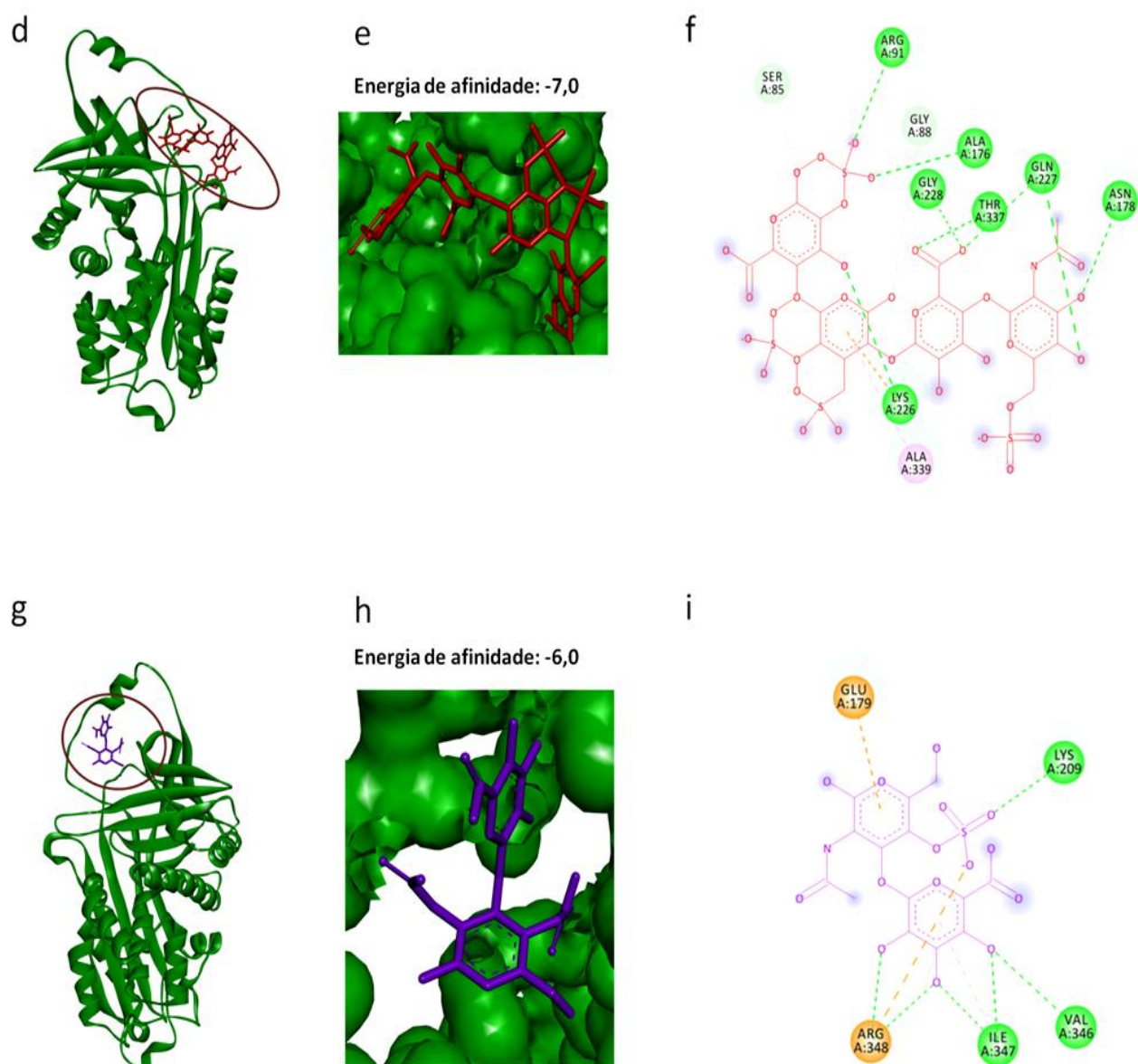


Figura 16. Docking entre TcSERPIN e a heparina de sódio (PubChem: 22833565) em três regiões diferentes: Em **a, d, g**, as diferentes regiões de ligação da proteína (em verde); em **b, e, h**, a região de cada ligação ampliada e mostrando as respectivas energias de ligação; em **c, f, i**, o mapa 2D de cada região de acoplamento, onde, o verde representa as ligações do tipo hidrogênio e amarelo e rosa, as ligações do tipo Pi.

5.11. Busca de alvo da rTcSERPIN em extrato de folhas do cacau

Para avaliar se o sistema de captura de proteases preparado com rTcSERPIN aprisionada a resina CNBr Sepharose interage com protease do extrato de folhas do cacau, uma análise em eletroforese em gel de zimograma foi realizada. Como resultado, a imagem do gel mostra um halo vertical claro ao lado do controle

positivo, a papaína de *Carica papaya* (16000 ng). É importante ressaltar que o gel de cobertura em gelatina foi confeccionado com tampão específico para papaína, contendo tampão fosfato, β -mercaptoetanol e EDTA. A visualização de um halo claro e contínuo ao longo do gel, demonstra que, possivelmente, mais de um tipo de protease da folha de *T. cacao* interagiu com rTcSERPIN (Figura 17).

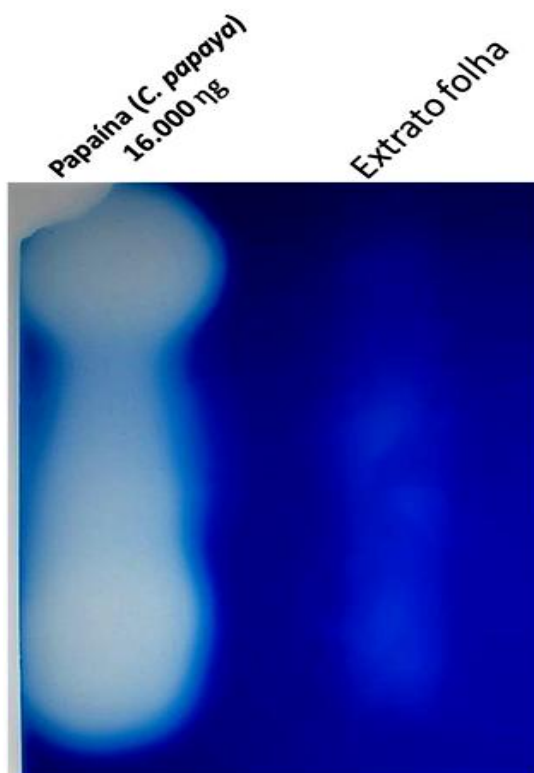


Figura 17. Análise em zimograma atividade de protease capturada de extratos de folhas de *T. cacao* com TcSERPIN imobilizada em resina CNBr-Sepharose: Linha 1, 16.000 ng de papaína de *Carica papaya* (controle positivo); linha 2, proteases capturadas por rTcSERPIN.

6. DISCUSSÃO

As análises *in silico* apresentaram uma proteína com 42,5 kDa, confirmada por análise de SDS-PAGE (Figura 9) após purificação. De maneira geral, serpinas são proteínas com tamanho médio, cujo peso varia em torno de 40 a 60 kDa graças a extensões das regiões N- e/ou C-terminais associados a elevados níveis de glicosilação (GETTINGS, 2002).

Entre as modificações pós-traducionais, a ausência de peptídeo sinal nas análises *in silico* surpreende, já que dados da literatura sobre serpinas de plantas, revelaram, em estudos *in vitro*, a presença de serpinas no apoplasto e em outros órgãos vegetais (LAMPL et al., 2010; VERCAMMEN et al., 2006), o que corrobora para o fato de que estes inibidores são secretados para outras rotas. A possível presença de TcSERPIN no espaço intercelular poderá ser confirmada por meio de captura utilizando proteases imobilizadas contra o fluido apoplástico, conforme método de extração de Pirovani et al (2008).

A análise da estrutura tridimensional de TcSERPIN (Figura 7) mostrou que a proteína possui 3 feixes de folhas- β , 9 α -hélices e a região do RCL conservada, demonstrada nos alinhamentos. Esse padrão estrutural típico da superfamília serpinas são fundamentais para sua estabilidade e a realização da atividade inibitória (STEIN & CARRELL, 1995; SILVERMAN et al., 2001). O RCL, principal sítio ativo das serpinas é responsável pela interação com a protease-alvo. Sua conservação estrutural é típico das serpinas inibitórias e possui em média 24 resíduos de aminoácidos (Figuras 4 e 5) (GETTINGS, 2002; ZHANG et al., 2017), próximos as folhas- β A, local onde o RCL clivado será inserido após interação com a protease alvo (GETTINGS, 2002; IRVING et al., 2000).

O tamanho do RCL pode influenciar nas taxa inibição das serpinas. A antitrombina, por exemplo, necessita de três resíduos a mais que a α_1 -antitripsina para conseguir a inserção do RCL no feixe de folhas- β , já a serpina de *Manduca sexta* com dois resíduos a menos que α_1 -antitripsina não consegue a inserção do RCL entre as folhas- β A (IRVING et al., 2000). Isso significa que o tamanho do RCL é um fator importante nos processos inibitórios das serpinas, já que é necessário haver uma inserção rápida do trecho do RCL clivado entre o feixe de folhas- β A para aprisionar a protease em uma ligação covalente com a serpina.

Outro fator importante para o sucesso da inserção do RCL entre as folhas- β está relacionado com a natureza dos resíduos presentes no principal sítio ativo do RCL (a região mais próxima à ocorrência da clivagem). Como observado nos dois alinhamentos (Figuras 4 e 6), as serpinas inibitórias mantêm seu padrão estrutural de conservação desde a glicina na posição 15 (região de dobradiça, “a *ringe region*”) à sequência de alaninas de p12~P9 (SILVERMAN et al., 2001; GETTINGS, 2002), esse ponto é crucial para o sucesso da ligação covalente exibido pelas serpinas e suas proteases cognatas. Serpinas que exibem um padrão diferente do descrito alteram seu potencial inibitório e perde a capacidade inibitória, tal como Maspin, mostrada no alinhamento (IRVING et al., 2000; GETTINGS, 2002). A serpina do cacau, TcSERPIN, exibe a região de sítio ativo do RCL “VNEEGTEAAA” (LAMPL et al 2010) (Figura 4) conservada, como descrita em *Arabidopsis*. Mutações nessa região, principalmente entre os resíduos P10 a P14 podem afetar a capacidade inibitória das serpinas (STEIN & CARRELL, 1995; CHEN et al., 2000).

Com base nos alinhamentos mostrados entre TcSERPIN e outras serpinas fica comprovado que o inibidor possui o sítio de clivagem conservado. A comparação de TcSERPIN com as duas serpinas de plantas como AtSerp1 de *A. thaliana* (LAMPL et al., 2010; VERCAMMEN et al., 2006) e a BSZx de *H. vulgare* (DAHL et al., 1996a, b), mostra que TcSERPIN possui similaridade em relação ao sítio de clivagem com suas homólogas. A principal região desse sítio P1 – P1’(ARG – GLY), local onde pode ocorrer a clivagem do RCL pela protease (CHEN et al., 2000) é idêntico ao de *A. thaliana*. O P1 da *H. vulgare* também é uma arginina, no entanto, a variação de P1’ nessa serpina é uma serina fazendo o par ARG/SER.

O RCL da maioria das serpinas em plantas exibe na região P1 uma arginina (Figura 6). Esse resíduo de aminoácido está presente em 79 % das serpinas do alinhamento. A região subsequente ao resíduo P1, o P1’ variou entre SER/GLY/ALA em todas as sequências. A conservação desses resíduos de aminoácidos na região do RCL, em plantas, dá a essas serpinas um padrão exclusivo presente somente em vegetais, quando comparado a outros organismos demonstrados no dendrograma (Figura 5). A análise desse dendrograma entre serpinas de plantas, humanos, insetos, termófilos, vírus e alguns vertebrados, apresentou uma distinta separação do clado *p* (clado de plantas) em relação aos outros grupos. Uma análise filogenética realizada por Irving e colaboradores (2000) já demonstrava que, evolutivamente, em algum momento da história, ocorreu uma divergência entre

serpinas ancestrais com posterior separação do clado *p*, relacionadas a sua estrutura e função. Esses dados sugerem que as serpinas devem ter aparecido no percurso histórico antes da diferenciação dos seres, ou seja, desde os primórdios da evolução (IRVING et al., 2000).

A análise filogenética também ganha suporte quando analisamos a Figura 4. Na comparação entre as serpinas de *Homo sapiens* com as três serpinas de plantas, é possível observar um padrão de conservação específico de vegetais, relacionados a manutenção de determinados resíduos de aminoácidos exclusivamente em plantas. Em 2010, Lampl e colaboradores chama a atenção para a manutenção de padrões de motivos conservados (Y XX G X D X R XF) e que são exclusivos em vegetais (LAMPL et al., 2010).

Os resultados de atividade de TcSERPIN *in vitro* demonstram que a proteína mantém sua atividade inibitória (Figura 12), embora tenha sido renaturada durante a diálise. O maior percentual inibitório de TcSERPIN foi obtido contra a papaína do mamão, com 54,6% em comparação aos 27,8% da tripsina suína. Aparentemente, o potencial inibitório de TcSERPIN, apesar de baixo, mostra que a proteína se manteve ativa após o re-enovelamento (*Refolding*). É preciso considerar também que a tripsina e a papaína utilizadas nos experimentos não são endógenas do cacau ou de um patógeno em co-evolução com o cacau, o que também pode explicar o baixo percentual inibitório. Também é importante ressaltar que devido ao alto poder metaestável desses inibidores, não sabemos até onde o *refolding* pode ter prejudicado sua atividade, já que serpinas são suscetíveis a mutações e perda de estabilidade (CHEN et al., 2000; GETTINGS, 2002). Porém, de modo seguro, os resultados obtidos mostram que TcSERPIN possui atividade inibitória *in vitro* contra proteases do tipo serina e cisteína.

Aparentemente, serpinas em plantas evoluíram para inibição de cisteíno-proteases. Isso pode ser devido a ausência de serino-proteases da família da quimotripsina relacionadas a essas serpinas, mas ausentes em plantas. AtSerp1 de *A. thaliana*, por exemplo, possui capacidade de inibição frente a cisteíno-proteases, como a metacaspase 9 (AtMC9) (VERCAMMEN et al., 2006), possivelmente, relacionada com morte celular programada da célula (VERCAMMEN et al., 2006; SRINIVASAN et al., 2009) e com uma papaína, a RD21, cujos testes *in vitro* e *in vivo* demonstraram que AtSerp1 inibiu essa protease e reduziu a morte celular na planta (LAMPL et al., 2010, 2013). Recentemente, um estudo com

nódulos de *Medicago truncatula* sob estresse hídrico mostrou níveis elevados de cisteíno-proteases, semelhantes a papaínas, e que uma serpina específica encontrada, nesse tecido, exerceu controle sobre essas proteases maximizando a sobrevivência da planta (DHANUSHKODI et al., 2018). De certa maneira, serpinas em plantas possui maior potencial inibitório para proteases do tipo cisteíno, em contraste com outros grupos de serpinas que são mais associadas com inibição frente a serino-protease da família da quimotripsina (OSTERGAARD et al., 2000).

Para analisar melhor a natureza das interações envolvendo serpinas foram realizadas algumas análises de *docking*, primeiro com duas serino-proteases com estruturas cristalografadas encontradas no banco de dados PDB, a α -quimotripsina e a α -tripsina pancreática bovinas e uma cisteíno-protease do látex do mamão.

Durante a interação entre a quimotripsina e TcSERPIN é possível visualizar a tríade catalítica (SER195, HIS57 e ASP103), em que a serina 195 da protease e a arginina 350, provável região P1 (região de clivagem segundo alinhamento) da serpina interagem por ligações de hidrogênio. No entanto, analisando a interação entre a serpina e a tripsina é possível visualizar algumas alterações, onde a serina 836 (posição correspondente a serina 195) da protease realiza ligações de hidrogênio com uma leucina 353 (região P3') da serpina, demonstrando uma possível variação do sítio de clivagem da serpina em comparação as duas serino-proteases (Figura 13).

A interação entre TcSERPIN e a cisteíno-protease do mamão também mostrou variação da interação dos resíduos do sítio de clivagem do RCL e o sítio ativo da protease. Na imagem é possível visualizar a cisteína 25 do sítio catalítico em ligação de hidrogênio com a valina 346 (posição P5) da serpina. Ainda, na imagem também é visível a presença da tríade (CYS25, HIS159 e ASN175) (Figura 13).

Esses resultados sugerem uma variação do sítio de clivagem do RCL entre a serpina do cacau e as três proteases cristalografadas. Foi demonstrado que a serpina de *Arabidopsis*, a AtSerp1, também possui variação do sítio de clivagem do RCL entre a posição P4-P8' (IKLRGLLMEEDE) (LAMPL et al., 2010). Ainda outro estudo analisando as regiões P1, P2 e P3 da Kallistatin, avaliou a especificidade desses resíduos como sítios de clivagem e demonstrou que mutações nessas posições por outros resíduos aumentavam ou reduziam a afinidade com as protease cognatas (CHEN et al., 2000). Esses resultados corroboram para o fato de que a

região de clivagem do RCL da serpina do cacau possa variar em relação a sua protease alvo, e, que, a extensão do trecho do RCL pode resultar em ligação peptídica entre a serpina e a protease ou pode transformar o inibidor em um simples substrato (WHISSTOCK & BOTTOMLEY, 2006; IRVING et al., 2000; GETTINS & OLSON, 2016).

Para esclarecer melhor a possível interação de TcSERPIN do cacau e possíveis alvos da via endógena, não somente foi preciso recorrer a dados da literatura (VERCAMMEN et al., 2006; LAMPL et al., 2010; DHANUSHKODI et al., 2018), como também foi necessário fazer uma rede de interação com base em sua homóloga de *Arabidopsis thaliana*, a AtSerp1. Com base nas rotas estabelecidas, a serpina do cacau pode estar relacionada em vias de estresse da planta, mais provavelmente, com proteases presente em vias de estresse e morte celular.

A via de interação de TcSERPIN com base em AtSerp1, mostra uma relação direta com duas cisteíno-proteases: a RD19 e a RD21, sendo que esta última já havia sido relacionada com AtSerp1, em que foi demonstrado interação de ambas as moléculas *in vitro* (LAMPL et al., 2010). Também é possível inferir que as serpinas do cacau e de *Arabidopsis* tenham papel regulatório diretamente com a RD21, já que, segundo a rede de interação, essa protease somente tem relação com a serpina no cluster selecionado. Com base nos resultados da rede e dados da literatura é que foram selecionadas homólogas da RD19 e da RD21 de *Arabidopsis* em *Theobroma cacao*, a partir do banco de dados do CocoaGenDB. As sequências obtidas foram modeladas e preparadas retirando o sítio pró-inibitório para a realização do *docking*.

As análises de *docking* com as cisteíno-proteases e a serpina do cacau retornou resultados distintos em relação aos resíduos do sítio de interação do RCL de TcSERPIN. A interação entre a RD21 e a serpina ocorreu de modo semelhante a papaína de *Carica papaya*, em que a valina 346 (P5) da serpina faz uma interação do tipo ligação de hidrogênio com a cisteína 167 (principal resíduo da tríade). Na imagem, ainda, é possível visualizar a presença da tríade catalítica da RD21 (CYS167, GLY161 e HIS303) (Figura 15b). No entanto, a interação de TcSERPIN e a RD19, a mostra a alanina 344 (região P7) da serpina com a cisteína 165 do sítio catalítico da protease. Na imagem a relação da tríade também está presente (CYS165, GLY159 e HIS308) (Figura 15d). Da mesma maneira como nas interações

das serino-proteases com a serpina do cacau, aparentemente, os resíduos do RCL de TcSERPIN variaram em cada interação.

A variação do sítio de clivagem do RCL pode implicar em aumento ou redução do potencial inibitório, pois a depender de onde ocorra a clivagem, o trecho do RCL poderia ficar muito curto ou longo demais para que houvesse rápida inserção entre as folhas- β A, ou seja, antes que a protease possa completar sua catálise. De certa maneira, o tamanho desse trecho e a afinidade dos resíduos do RCL com as proteases cognatas pode aumentar ou reduzir o potencial de inibição, significando que TcSERPIN possa ter um papel regulatório, seja como inibidor ou substrato (GETTINS & OLSON, 2016).

Juntando esses dados com a análise do eluído da folha relacionado a captura da protease pela serpina acoplada a CNBr *Sepharose*, podemos inferir que TcSERPIN interage com cisteíno-proteases, mais precisamente, papaínas, já que a confecção do gel foi realizada com tampão fosfato, EDTA e β -mercaptoetanol, situação mais favorável para essas proteases. Embora ainda se conheça pouco sobre a possível relação das serpinas e outras vias dentro da célula, o fato de que possa haver essa interação entre TcSERPIN e papaínas do cacau é importante, pois cisteíno-proteases do tipo papaínas podem estar relacionadas em vias de defesa de plantas, com respostas à estresse e morte celular programada (SCHALLER, 2004; LI et al., 2008). E isso pode significar que serpinas em plantas possam ter um papel regulatório relacionado a essas vias.

Cisteíno-proteases como a RD21 parecem estar presentes em estado de senescência e estresse da planta (KINOSHITA et al., 1999). Outros inibidores de cisteína, tais como cistatina, por exemplo, pode atrasar a morte celular induzida por estresse (LAMPL et al., 2010), mediar defesa contra patógenos como *Moniliophthora perniciosa*, e estar presente em rotas de necrose e morte celular programada (PIROVANI et al., 2010; CARDOSO et al., 2015); em outro exemplo, as metacaspases, outro tipo de cisteíno-protease, podem ativar a morte celular programada em leveduras e em plantas (WATANABE & LAM, 2005; BOSCH & FRANKLIN-TONG, 2007; HE et al., 2008). No entanto, ainda que se saiba que essas cisteíno-proteases possam estar relacionadas a processos de defesa e estresse da planta, ainda não se conhece ao certo como exercem seu controle, nem mesmo como podem estar relacionadas com serpinas (LAMPL et al., 2010, 2013).

Outra análise de *docking* realizada levou em consideração comparações entre a hélice D de TcSERPIN e a hélice D da Antitrombina III em relação a heparina de sódio. Os sítios selecionados para o *docking* de TcSERPIN com a heparina levaram em consideração a relação de afinidade da hélice D da antitrombina III até o RCL. Ocorreu ligação da heparina nos sítios selecionados com baixa energia de ligação, ou seja, alta afinidade nas regiões testadas (Figura 16). Este resultado se torna de fundamental importância para nossos estudos, pois quando a protease reconhece a região P1 (ARG) da Antitrombina III, nesse momento, o RCL é parcialmente inserido entre as folhas- β A. Essa interação se configura por ser relativamente fraca e lenta. No entanto, a ligação da heparina no sítio que corresponde à hélice D da serpina estabelece uma junção entre as duas proteínas (protease e inibidor), o que mantém a protease aprisionada à serpina e faz com que haja uma reorientação da estrutura do inibidor, potencializando em muitas vezes as taxas de inibição (WHISSTOCK et al., 2000; OLSON, 2010; GETTINS & OLSON, 2009, 2016).

A afinidade da heparina de sódio com TcSERPIN poderá ser testado *in vitro* para sabermos se essa ligação pode efetivar o potencial inibitório da serpina do cacau da mesma maneira que na Antitrombina III de humanos. Além disso, ainda não foi demonstrado se as serpinas de plantas possuem um carboidrato endógeno ou do fitopatógeno que possam desempenhar função similar à heparina, como facilitador da interação entre o inibidor e a protease alvo.

Por fim, a análise por dicroísmo circular, *unfolding* e *refolding*, mostrou que, apesar de rTcSERPIN ter sido desnaturada e renaturada com a retirada gradativa da uréia, segundo os resultados dos testes de inibição, a proteína permaneceu ativa. No entanto, quando submetida a aumentos de temperatura, acima de 35° e 66° C em pH 5,4 e pH 7,0, respectivamente, a proteína perdeu sua estrutura secundária e precipitou. E durante o *refolding*, TcSERPIN não retornou a estrutura inicial. Apesar desse resultado surpreender, no entanto, não é novo. Um estudo com uma serpina de *Thermobifida fusca*, a Thermopin, de microrganismos termófilos, resistente a altas temperaturas e com Tm de 65 °C, também formou agregados com perda de solubilidade e atividade quando submetida a altas temperaturas. Outros resultados similares também foram mencionados para Antitrombina e Maspin (IRVING et al., 2003). Aparentemente, serpinas são propensas à formação de polímeros (IRVING et al., 2002).

7. CONCLUSÃO

Os resultados deste trabalho demonstram que TcSERPIN possui conservação dos seus resíduos de aminoácidos e que a sequência primária do RCL, principal local de interação da proteína com as peptidases, possui o mesmo padrão de outras proteínas homólogas e com potencial inibitório.

A purificação da rTcSERPIN a partir da fração insolúvel do extrato de *E. coli* apresentou grande rendimento e resultou em uma proteína ativa mesmo após a sua desnaturação por uréia. Embora a proteína não re-enovele após o *Unfolding* por elevação de temperatura durante as análises por dicroísmo circular (CD), a rTcSERPIN mostrou o mesmo padrão de outras serpinas, com predisposição para formação de aglomerados, sob variações de pH e temperatura, alterando sua estabilidade, e, possivelmente, sua conformação tridimensional.

Os testes de inibição e *docking* comprovam que a proteína interage com peptidases do tipo serino e cisteíno, porém, com maior afinidade com papaínas. Além disso, a rTcSERPIN também demonstra interagir com cisteíno-proteases de cacau, segundo a análise em zimograma.

As análises até o momento comparam os resultados obtidos até aqui com outros resultados da literatura e demonstram que a proteína TcSERPIN pode ser relacionada com vias de regulação e morte celular programada. Esses resultados são de grande utilidade e podem significar que a proteína recombinante tem potencial para ser utilizada em testes futuros, visando a resistência da planta durante a interação planta/patógeno.

8. REFERÊNCIAS

ABBENANTE, G.; FAIRLIE, D. P. Protease inhibitors in the clinic. **Medicinal Chemistry**, v. 1, p. 71-104, 2005.

AHN, J. W.; ATWELL, B. J.; ROBERTS, T. H. Serpin genes AtSRP2 and AtSRP3 are required for normal growth sensitivity to a DNA alkylating agent in *Arabidopsis*. **BMC Plant Biology**, v. 9, n. 52, p. 1 - 11, 2009.

AIME, M. C.; PHILLIPS-MORA, W. The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. **Mycologia**, v. 97, p. 1012-1022, 2005.

ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids**, v. 25, n. 17, p. 3389-3402, 1997.

ALVAREZ-ALFAGEME, F. *et al.* Potential Use of a Serpin from *Arabidopsis* for Pest Control. **Plos One**, v. 6, n. 5, p. 1-9, 2011.

ARGOUT, X. RUIZ, M.; TURNBULL, C. LANAUD, C.; ROSENQUIST, E. CocoaGen DB: a Web portal for crossing cocoa phenotypic, genetic and genomic data from ICGD and TropGeneDB data bases. In: **15th International Cocoa Research Conference**, v. 1, p. 512-518, San Jose, Costa Rica, 2007.

ARGOUT, X.; FOUET, O.; WINCKER, P.; GRAMACHO, K.; LEGAVRE, T.; SABAU, X.; RISTERUCCI, A. M.; DA SILVA, C.; CASCARDO, J.; *et al.* Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao* L. generated from various tissues and under various conditions. **BMC Genomics**, v. 9, p. 548, 2008.

ARGOUT, X.; SALSE, J.; AURY, J. M.; GUILTINAN, M. J.; DROC, G.; GOUZY, J.; ALLEGRE, M.; CHAPARRO, C.; LEGAVRE, T.; MAXIMOVA, S. N.; ABROUK, M.; MURAT, F.; *et al.* *Nature Genetics*, v. 42, n. 2, p. 101-118, 2011.

ARNOLD, K.; BORDOLI, L.; KOPP, J.; SCHWEDE, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. **Bioinformatics**, v. 22, n. 2, p. 195-201, 2006.

BARBOSA, C. S.; FONSECA, R. R. D.; BATISTA, T. M.; BARRETO, M. A.; ARGOLLO, C. S.; CARVALHO, M. R.; AMARAL, D. O. J. D.; SILVA, E. M. A.; ARÉVALO-GARDINI, E.; HIDALGO, K. S.; FRANCO, G. R.; PIROVANI, C. P.; MICHELI, F.; GRAMACHO, K. P. Genome sequence and effectorome of *Moniliophthora perniciosa* and *Moniliophthora roreri* subpopulations. **BMC Genomics**, v. 19, n.1, p. 509, 2018.

BERG, J. M.; TYMOCZKO, J. L.; STRYER, L. **Bioquímica**. 6 ed. Rio de Janeiro: Guanabara, 2008. 1114 p.

BHATTACHARJEE, L.; SINGH, P. K.; SINGH, S.; NANDI, A. K. Down-regulation of rice serpin gene OsSRP-LRS exaggerates stress-induced cell death. **Journal of Plant Biology**, v. 58, n. 5, p. 327 - 332, 2015.

BHATTACHARJEE, L.; SINGH, D. K.; GAUTAM, J. K.; NANDI, A. K. Arabidopsis thaliana serpins AtSRP4 and AtSRP5 negatively regulate stress-induced cell death and effector-triggered immunity induced by bacterial effector AvrRpt2. **Physiologia Plantarum**, v. 159, n. 3, p. 329 - 339, 2017.

BLOM, N.; GAMMELTOFT, S.; BRUNAK, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. **Journal of Molecular Biology**, v. 294, n. 5, p. 1351-1362, 1999.

BOSCH, M.; FRANKLIN-TONG, V. E. Temporal and spatial activation of caspase-like enzymes induced by self-incompatibility in Papaver pollen. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 46, p. 18327-18332, 2007.

BULLER, A. R.; TOWNSEND, C. A. Intrinsic evolutionary constraints on protease structure, enzyme acylation, and the identity of the catalytic triad. **PNAS Plus**, v. 110, n. 8, E653–E661. 2013.

CARDOSO, T. H.; FREITAS, A. C.; ANDRADE, B. S.; SOUSA, A. O.; SANTIAGO ADA, S.; KOOP, D. M.; GRAMACHO, K. P.; ALVIM, F. C.; MICHELI, F.; PIROVANI, C. P. TcCYPR04, a Cacao papain-Like cysteine-protease detected in senescent and necrotic tissues interacts with a cystatin TcCYS4. **PLoS One**, v. 10, n. 12, e0144440, 2015.

CARRELL, R. W.; TRAVIS, J. α 1-Antitrypsin and the serpins. Variation and countervariation. **Trends in Biological Sciences**, v. 10, p. 20–24, 1985.

CARRELL, R. W.; STEIN, P.E.; FERMI, G.; WARDELL, M. R. Biological implications of a 3 Å structure of dimeric antithrombin. **Structure**, v. 2, p. 257–270, 1994.

CHEN, V. C.; CHAO, L.; CHAO, J. Papéis dos resíduos P1, P2 e P3 na determinação da especificidade inibitória da calistatina em relação à calicreína tecidual. **The Journal of Biological Chemistry**, v. 275, p. 38457-38466, 2000.

CHEN, H.; ZHENG, D.; AMBADAPADI, S.; DAVIDS, J.; RYDEN, S.; SAMY, H.; BARTEE, M.; SOBEL, E.; DAI, E.; LIU, L.; MACAULAY, C.; YACHNIS, A.; WEYAND, C.; THOBURN, R.; LUCAS, A. Serpin Treatment Suppresses Inflammatory Vascular Lesions in Temporal Artery Implants (TAI) from Patients with Giant Cell Arteritis. **PloS One**, v. 10, N. 2, n. 4, p. 1-15, 2015.

CHEN, H.; ZHENG, D.; DAVIDS, J.; BARTEE, M. Y.; DAI, E.; LIU, L.; PETROV, L.; MACAULAY, C.; THOBURN, R.; SOBEL, E.; MOYER, R.; MCFADDEN, G.; LUCAS, A. VIRAL Serpin Therapeutics: From Concept to Clinic. **Methods Enzymology**, v. 499, p. 301-329, 2013.

CRAIK, C.; PAGE, M. MADISON, E. Proteases as therapeutics. **Biochemical Journal**, v. 435, p. 1-17, 2011.

DAHL, S. W.; RASMUSSEN, S. K.; HEJGAARD, J. Heterologous expression of three plant serpins with distinct inhibitory specificities. **Journal of Biological Chemistry**, v. 271, p. 25083–25088, 1996a.

DAHL, S. W.; RASMUSSEN, S. K.; PETERSEN, L. C.; HEJGAARD, J. Inhibition of coagulation factors by recombinant barley serpin BSZx. **FEBS Letters**, v. 394, n. 2, p. 165 – 168, 1996b.

DHANUSHKODI, R.; MATTHEW, C.; MCMANUS, M. T.; DIJKWEL, P. P. Drought-induced senescence of *Medicago truncatula* nodules involves serpin and ferritin to control proteolytic activity and iron levels. **New Phytologist**, v. 220, n. 1, p.196–208, 2018.

DZNIC, S. H.; BERNARDO, M. M.; LI, X.; FERNANDEZ-VALDIVIA, R.; HO, Y. S.; MI, Q. S.; BANDYOPADHYAY, S.; LONARDO, F.; VRANIC, S.; OLIVEIRA, D. S.; BONFIL, R. D.; DYSON, G.; CHEN, K.; OMEROVIC, A.; SHENG, X.; HAN, X.; WU, D.; BI, X.; CABARAVDIC, D.; JAKUPOVIC, U.; WAHBA, M.; PANG, A.; HARAJLI, D.; SAKR, W. A.; SHENG, S. An essential Role of Maspin in Embryogenesis and Tumor Suppression. **Cancer Research**, v. 77, n. 4, p. 886-896, 2017.

ELLIOT, P. R., ABRAHAM, J. P.; LOMAS, D. A. Wild-type alpha 1-antitrypsin is in the canonical inhibitory conformation. **Journal of Molecular Biology**, v. 275, n. 3, p. 419-425 1998.

EMANUELSSON, O.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. Locating proteins in the cell using TargetP, SignalP and related tools. **Nat Protocols**, v. 2, n. 4, p. 953-971, 2007.

EVANS, H. C.; BEZERRA, J. L.; BARRETO, R. W. Of mushrooms and chocolate trees: aetiology and phylogeny of witches' broom and frosty pod diseases of cacao. **Plant Pathology**, v. 62, n. 4, p. 728–740, 2013.

FUJINAGA, M.; CHERNEY, M. M.; OYAMA, H.; ODA, K.; JAMES, M.N. The molecular structure and catalytic mechanism of a novel carboxyl peptidase from *Scytalidium lignicolum*. **Proceedings of the National Academy of Sciences**, v. 101, p. 3364–3369, 2004.

GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; DUVAUD, S.; WILKINS, M. R.; APPEL, R. D.; BAIRDOCH, A. Protein Identification and Analysis Tools on the ExPASy Server. In: **The Proteomics Protocols Handbook**. Totowa, NJ: Humana Press, 2005. p. 571-607.

GESTEIRA, A.; MICHELI, F.; CARELS, N.; SILVA, A. C.; GRAMACHO, K. P.; SCHUSTER, L.; MACÊDO, J. N.; PEREIRA, G. A. G.; CASCARDO, J. C. M. Comparative Analysis of Expressed Genes from Cacao Meristems Infected by *Moniliophthora perniciosa*. **Annals of Botany**, v. 100, p. 129-140, 2007.

GETTINS, P. G. Serpin structure, mechanism, and function. **Chemical Reviews**, v. 102, n. 12, p. 4751-4804, 2002.

GETTINS, P. G. W.; OLSON, S. T. Activation of antithrombin as a factor IXa and Xa inhibitor involves mitigation of repression rather than positive enhancement. **FEBS Letters**, v. 583, n. 21, p. 1-9, 2009.

GETTINS, P. G. W.; OLSON, S. T. Inhibitory Serpins. New Insights into their Folding, Polymerization, Regulation and Clearance. **Biochemical Journal**, v. 473, n. 15, p. 2273–2293, 2016.

GRIFFITH, W. G.; NICHOLSON, J.; NENNINGER, A.; IRCH, R. M.; HEDGER, J. N. Witches' brooms and frosty pods: two major pathogens of cacao. **New Zealand Journal of Botany**, v. 41, p. 423-435, 2003.

GUPTA, R.; JUNG, E.; BRUNAK, S. Prediction of N-glycosylation sites in human proteins. *Pacific Symp.* **Biocomput**, v. 7, p. 310-322, 2004

HAMERSCHLAK, N.; ROSENFELD, L. G. M. Utilização da Heparina e dos Anticoagulantes Orais na Prevenção e Tratamento da Trombose Venosa Profunda e da Embolia Pulmonar. **Arquivos Brasileiros de Cardiologia**, v. 67, n. 3, p. 209-213, 1996.

HAQ S. K.; ATIF S. M.; KHAN R. H. Protein proteinase inhibitor genes in combat against insects, pests, and pathogens: natural and engineered phytoprotection. **Archives of Biochemistry and Biophysics**, v. 431, p.145-159, 2004.

HE, R.; DRURY, G. E.; ROTARI, V. I.; GORDON, A.; WILLER, M.; FARZANEH, T.; WOLTERING, E. J.; GALLOIS, P. Metacaspase-8 modulates programmed cell death induced by ultraviolet light and H₂O₂ in Arabidopsis. **The Journal of Biological Chemistry**, v. 283, n. 2, p. 774–783, 2008.

HEIT, C.; JACKSON, B. C.; MACANDREWS, M.; Wright, M. W.; Thompson, D.C.; Silverman, G. A.; Nebert, D. W.; Vasiliou, V. Update of the human and mouse SERPIN gene superfamily. **Human Genomics**, v. 7, p. 1-22, 2013.

HOPPINKS, P. C.; CARREL, R. W.; STONE, S. R. Effects of mutations in the hinge region of serpins. **Biochemistry**, v. 32, p. 7650 – 7657, 1993.

HOPPINKS, P. C.; CARREL, R. W.; STONE, S. R. Effects of mutations in the hinge region of serpins. **Biochemistry**, v. 32, p. 7650 – 7657, 1993.

HOU, S.; JAMIESON, P.; HE, P. The cloak, dagger, and shield: proteases in plant–pathogen interactions. **Biochemical Journal**, v. 475, p. 2491–2509, 2018.

HUNT, L. T.; DAYHOFF, M. O. A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha 1-proteinase inhibitor. **Biochemical Biophysical Research Commun**, v. 95, n. 2, p. 864-871, 1980.

- HUNTINGTON, J. A.; READ, R. J.; CARRELL, R. W. Structure of a serpin-protease complex shows inhibition by deformation. **Nature**, v. 407, n. 6806, p. 923-926, 2000.
- IRVING, J. A.; PIKE, R. N.; LESK, A. M.; WHISSTOCK, J. C. Phylogeny of the serpin superfamily: Implications of patterns of amino acid conservation for structure and function. **Genome Research**, v. 10, p.1845-1864, 2000.
- IRVING, J. A.; STEENBAKKERS, P. J.; LESK, A. M.; OP DEN CAMP, H. J.; PIKE, R. N.; WHISSTOCK, J. C. Serpins in prokaryotes. **Molecular Biology and Evolution**, v. 19, n. 11, p. 1881-90, 2002.
- IRVING, J. A.; CABRITA, L. D.; ROSSJOHN, J.; PIKE, R. N.; BOTTOMLEY, S. P.; WHISSTOCK, J. C. The 1.5 Å Crystal Structure of a Prokaryote Serpin. **Structure**, v. 11, n. 4, p. 387–397, 2003.
- JIN, L.; ABRAHAMS, J. P.; SKINNER, R.; PETITOU, M.; PIKE, R. N.; CARRELL, R. W. The anticoagulant activation of antithrombin by heparin. **Proceedings of the National Academy of Sciences U.S.A.**, v. 94, n. 26, n. 14683–14688, 1997.
- JUDD, W. S.; CAMPBELL, C. S.; KELLOGG, E. A.; STEVENS, P. F.; DONOGHUE, M. J. **Sistemática vegetal: Um enfoque filogenético**. 3. ed. Porto Alegre: Artmed, 2009. 632p.
- KHAN, M. S.; SINGH, P.; AZHAR, A.; NASEEM, A.; RASHID, Q.; KABIR, M. A.; JAIRAJPURI, M. A. Serpin Inhibitor Mechanism: A Delicate Balance between Native Metastable State and Polymerization. **Journal of Amino Acids**, v. 2011, p. 1-10, 2011.
- KIEMER, L.; BENDTSEN, J. D.; BOLM, N. **NetAcet**: Prediction of N-terminal acetylation sites. **Bioinformatics**, v. 21, n. 7, p. 1269-1270, 2004.
- KINOSHITA, T.; YAMADA, K. HIRAIWA, N.; KONDO, M.; NISHIMURA, M.; HARA-NISHIMURA, I. Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions. **The Plant Journal**, v. 19, n. 1, p. 43-53, 1999.
- KOZAKOV, D.; HALL, D. R.; XIA, B.; PORTER, K. A.; PADHORN, D.; YUEH, C.; BEGLOV, D.; VAJDA, S. The ClusPro web server for protein-protein docking. **Nature Protocols**, v.12, n. 2, p. 255-278, 2017.
- KRAUT, J. Serine-proteases - structure and mechanism of catalysis. **Annual Review of Biochemistry**, v. 46, p. 331-358, 1977.
- KROWARSCH, D.; CIERPICKI, T.; JELEN, F.; OTLEWSKI, J. Canonical protein inhibitors of serine proteases. Review. **Cellular and Molecular Life Sciences**, v. 60, p. 2427– 2444, 2003.
- LAMPL, N.; HADRIAN-BUDAI, O.; DAVYDOV, O.; JOSS, T. V.; HARROP, S. J.; CURMI, P. M. G.; ROBERTS, T. H.; FLUHR, R. *Arabidopsis* AtSerp1, crystal structure and *in vivo* interaction with its target protease RESPONSIVE TO

DESICCATION-21 (RD21). **The Journal of Biological Chemistry**, v. 285, n. 18, p. 13550-13560, 2010.

LAMPL, N.; ALKAN, N.; DAVYDOV, O.; FLUHR, R. Set-point control of RD21 protease activity by AtSerp1 controls cells death in *Arabidopsis*. **The Plant Journal**, v. 74, n. 3, p. 498-510, 2013.

LASKOWSKI, R. A.; MACARTHUR, M. W.; MOSS, D. S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283-291, 1993.

LAU, J. K-C.; CHENG, Y-K. SER-HIS-ASP catalytic triad in model non-aqueous solvent environment: A computational study. **Bioorganic & Medicinal Chemistry Letters**, v. 16, n. 22, p. 5797–5800, 2006.

LEHNINGER, A. L.; NELSON, D. L.; COX, M. M. **Princípios de Bioquímica**. 2. ed. São Paulo: Sarvier, 2000. 839 p.

LEMA ASQUI, S.; VERCAMMEN, D.; SERRANO, I.; VALLS, M.; RIVAS, S.; VAN BREUSEGEM, F.; CONLON, F. L.; DANGL, J. L.; COLL, N. S. AtSerp1 is an inhibitor of the metacaspase AtMC1-mediated cell death and autocatalytic processing in planta . **New Phytologist**, v. 218, n. 3, p. 1156–1166, 2017.

LI, J.; BRADER, G.; PALVA, E. T. Kunitz Trypsin Inhibitor: An Antagonist of Cell Death Triggered by Phytopathogens and Fumonisin B1 in *Arabidopsis*. **Molecular Plant**, v. 1, n. 3, p. 482-495, 2008.

LI, X.; JIN, Z.; GAO, F.; LU, J.; CAI, L.; DONG, J.; YU, J.; YANG, H . Characterization of barley serpin Z7 that plays multiple roles in malt and beer. **Journal of Agricultural and Food Chemistry**, v. 62, n. 24, p. 5643-5650, 2014.

LOEBERMANN, H.; TOKUOKA, R.; DEISENHOFER, J.; HUBER, R. Human α 1-proteinase inhibitor. **Journal of Molecular Biology**, v. 177, p. 531-556, 1984.

LORENZI, H.; MATOS, F. J. A. **Plantas Medicinais no Brasil: Nativas e Exóticas**. 2. ed. São Paulo: Nova Odessa, 2002. 512 p.

LOUIS-JEUNE, C.; ANDRADE-NAVARRO, M. A.; PEREZ-IRATXETA, C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. **Proteins**, v. 80, n. 2, p. 374-381, 2012.

LUCAS, A.; YARON, J. R.; ZHANG, L.; MACAULAY, C.; MCFADDEN, G. Serpins: Development for Therapeutic Applications. **Serpins: Methods and Protocols**, v. 1826, p. 255–265, 2018.

LUENGO-GI, G.1.; CALVO, M. I.; MARTÍN-VILLAR, E.; ÁGUILA, S.; BOHDAN, N.; ANTÓN, A. I.; ESPÍN, S.; AYALA DE LA PEÑA, F.; VICENTE, V.; CORRAL, J.; QUINTANILLA, M.; MARTÍNEZ-MARTÍNEZ, I. Antithrombin controls tumor migration, invasion and angiogenesis by inhibition of enteropeptidase. **Scientific report**, v. 6, p. 27544, 2016.

MELO, F.; FEYTMANS, E. Assessing protein structures with a non-local atomic interaction energy. **Journal of Molecular Biology**, v. 277, n. 5, p. 1141-1152, 1998.

MICHAUD, D.; CANTIN, L.; RAWORTH, D. A.; VRAIN, T. C. Assessing the stability of cystatin/cysteine proteinase complexes using mildlydenaturing gelatin-polyacrylamide gel electrophoresis. **Electrophoresis**, v. 17, n. 1, p. 74–79, 1996.

MONDEGO, J. M.; CARAZZOLLE, M. F.; COSTA, G. G.; FORMIGHIERI, E. F.; PARIZZI, L. P.; RINCONES, J.; COTOMACCI, C.; CARRARO, D. M.; CUNHA, A. F.; CARRER, H.; et al. A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao. **BMC Genomics**, v. 18, n. 9, p. 548, 2008.

NAKAI, K.; HORTON, P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. **Trends in Biochemical Sciences**, v. 24, n. 1, p. 34-35, 1999.

OLSON, S. T.; BJÖRK, I. Predominant contribution of surface approximation to the mechanism of heparin acceleration of the antithrombin-thrombin reaction. Elucidation from salt concentration effects. **The Journal of Biological Chemistry**, v. 266, n. 10, p. 6353–6364, 1991.

OLSON, S. T.; RICHARD, B.; IZAGUIRRE, G.; SCHEDIN-WEISS, S. GETTINS, P. G. Molecular mechanisms of antithrombin-heparin regulation of blood clotting proteinases. A paradigm for understanding proteinase regulation by serpin family protein proteinase inhibitors. **Biochimie**, v. 92, n. 11, p. 1587-1596, 2010.

OSTERGAARD, H.; RASMUSSEN, S. K.; ROBERTS, T. H.; HEJGAARD, J. Inhibitory serpins from wheat grain with reactive centers resembling glutamine-rich repeats of prolamin storage proteins. Cloning and characterization of five major molecular forms. **The Journal of Biological Chemistry**, v. 275, n. 43, p. 33272-33279, 2000

OTLEWSKI, J.; JELEN, F.; ZAKRZEWSKA, M.; OLESY, A. The many faces of protease-protein inhibitor Interaction. **The EMBO Journal**, v. 24, p.1303-1310, 2005.

PATSTON, P. A.; CHURCH, F. C.; OLSON, S. T. "Serpín-ligand interactions." **Methods**, v. 32, n. 2, p. 93-109, 2004.

PETERSEN, T. N.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, n. 10, p. 785-786, 2011.

PIROVANI, C. P.; MACEDO, J. N.; CONTIM, L. A.; MATRANGOLO, F. S.; LOUREIRO, M. E.; FONTES, E. P. A sucrose-binding protein homologue from soybean exhibits GTP-binding activity that functions independently of sucrose transport activity. **European Journal of Biochemistry**, v. 269, n. 16, p. 3998–4008, 2002.

PIROVANI, C. P.; SANTIAGO, A. S.; SANTOS, L. S.; MICHELI, F.; MARGIS, R.; GESTERIA, A. S.; ALVIM, F. C.; PEREIRA, G. A.; DE MATTOS CASCARDO, J. C. *Theobroma cacao* cystatins impair *Moniliophthora perniciosa* mycelial growth and are involved in postponing cell death symptoms. **Planta**, v. 232, n. 6, p. 1485–1497, 2010.

PIROVANI, C. P.; CARVALHO, H. A.; MACHADO, R. C.; GOMES, D. S.; ALVIM, F. C.; POMELLA, A. W. et al. Protein extraction for proteome analysis from cacao leaves and meristems, organs infected by *Moniliophthora perniciosa*, the causal agent of the witches' broom disease. **Electrophoresis**, v. 29, p. 2391–2401, 2008.

PLOETZ, R. C. Cacao Diseases: Important Threats to Chocolate Production Worldwide. **Phytopathology**, v. 97, n. 12, p. 1634–1639, 2007.

RASMUSSEN, S. K.; DAHL, S. W.; NORGARD, A.; HEJGAARD, J. A recombinant wheat serpin with inhibitory activity. **Plant Molecular Biology**, v. 30:673–677, 1996.

RAWLINGS, N. D.; BARRETT, A. J.; BATEMAN, A. Asparagine Peptide Lyases a seventh catalytic type of proteolytic enzymes. **The journal of Biological Chemistry**, v.12, p.38321-38328, 2011.

RAWLINGS, N. D.; BARRETT, A. J.; THOMAS, P. D.; HUANG, X.; BATEMAN, A.; FINN, R. D. The MEROPS: the database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. **Nucleic Acids Research**, v. 46, p. D624–D632, 2018.

RAWLINGS, N. D.; TOLLE, D. P.; BARRET, A. J. Evolutionary families of peptidase inhibitors. **Biochemistry Journal**, v. 378, p. 705-716, 2004.

RAWLINGS N.D.; BARRETT J.; Evolutionary families of peptidases. **The Biochemical Journal**, v. 290, p. 205–218, 1993.

RAWLINGS, N. D.; BARRETT, A. J. MEROPS: the peptidase database. **Nucleic Acids Research**, v. 27, p. 325–331, 1999.

RAWLINGS N. D.; BARRETT A. J.; BATEMAN A. Asparagine Peptide Lyases a seventh catalytic type of proteolytic enzymes. **The journal of Biological Chemistry**, v. 286, n. 44, p. 38321–38328, 2011.

ROBERTS, T. H.; HEJGAARD, J. Serpins in plants and green algae. **Function and Integrative Genomics**, v. 8, p.1-27, 2008.

RUSSO, C. **Modelagem de serino-proteases e inibidores com emprego de ferramentas de bioinformática estrutural**. 2006. 79 f. Dissertação (Mestrado em Biologia Celular e Molecular) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

RYAN, C. A. Protease inhibitors in plants: genes for improving defense against insects and pathogens. **Annual Review of Phytopathology**, v. 28, p. 425–449, 1990.

SANNER, M.; OLSON, A. J.; SPEHNER, J. C. Reduced Surface: an Efficient Way to Compute Molecular Surfaces. **Biopolymers**, v. 38, n. 3, p. 305-320, 1996.

SANTOS, A. F. *et al.* Peptidases em biotecnologia: produção, aplicações e mercado. In.: **Biotecnologia aplicada a agroindústria: fundamentos e aplicações**. v. 4. São Paulo: Blucher, 2016. Cap. 11, p. 381-437.

SCHALLER, A. A cut above the rest: the regulatory function of plant proteases. **Planta**, v. 220, n. 2, p. 183-197, 2004.

SEEMULLER E.; LUPAS A.; STOCK D.; LOWE J.; HUBER R.; BAUMEISTER W. Proteasome from *Thermoplasma acidophilum*: A Threonine Protease. **SCIENCE**, v. 268, p. 579-582, 1995.

SHINDO, T.; VAN DER HOORN, R. A. L. Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. **Molecular Plant Pathology**, v. 9, n. 1, p. 119-125, 2008.

SILVERMAN, G. A.; BIRD, P. I.; CARRELL, R. W.; CHURCH, F. C.; COUGHLIN, P. B.; GETTINS, P. G.; IRVING, J. A.; LOMAS, D. A.; LUKE, C. J.; MOYER, R. W.; PEMBERTON, P. A.; REMOLD-O'DONNELL, E.; SALVESEN, G. S.; TRAVIS, J.; WHISSTOCK, J. C. The serpins are an expanding superfamily of structurally similar but functionally diverse proteins: evolution, mechanism of inhibition, novel functions, and a revised nomenclature. **Journal of Biological Chemistry**, v. 276, p. 33293-33296, 2001.

SOUZA, V.C.; LORENZI, H. 2012. **Botânica Sistemática: guia ilustrado para identificação das famílias de Fanerógamas nativas e exóticas no Brasil, baseado em APG III**. 3ª ed. Nova Odessa, São Paulo: Instituto Plantarum, 2012. 768p.

SRINIVASAN, T.; KUMAR, K. R.; KIRTI, P. B. Constitutive expression of a trypsin protease inhibitor confers multiple stress tolerance in transgenic tobacco. **Plant Cell Physiology**, v. 50, p. 541–553, 2009

STEIN, P. E.; CARRELL, R. W. What do dysfunctional serpins tell us about molecular mobility and disease? **Nature Structural & Molecular Biology**, v. 2, p. 96-113, 1995.

TANAKA, N.; FUJITA, M.; HANDA, H.; MURAYAMA, S.; UEMURA, M.; KAWAMURA, Y.; MITSUI, T.; MIKAMI, S.; TOZAWA, Y.; YOSHINAGA, T.; KOMATSU, S. Proteomics of the rice cell: systematic identification of the protein populations in subcellular compartments. **Molecular Genetics and Genomics**, v. 271, n. 5, p. 566–576, 2004.

TIPTON, K. F. Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB). Enzyme nomenclature. Recommendations.

Supplement: corrections and additions. **European Journal of Biochemistry**, v. 223, n. 1, p. 1-5, 1992.

TORUÑO, T.Y., STERGIOPOULOS, I.; COAKER, G. Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. **Phytopathology**, v. 54, p. 419–441, 2016.

TROTT, O.; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. **Journal of Computational Chemistry**, v. 31, n. 2, p. 455-461, 2010.

VAN BOVEN, H H.; LANE, D. A. Review Antithrombin and its inherited deficiency states. **Seminars Hematology**, v. 34, n.3, p. 188-204, 1997.

VERCAMMEN, D.; BELENGHI, B.; VAN DE COTTE, B.; BEUNENS, T.; GAVIGAN, J. A.; DE RYCKE, R.; BRACKENIER, A.; INZE, D.; HARRIS, J. L.; VAN BREUSEGEM, F. Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for Metacaspase 9. **Journal Molecular Biology**, v. 364, p. 625–636, 2006.

WATANABE, N.; LAM, E. Two *Arabidopsis* metacaspases AtMCP1b and AtMCP2b are arginine/lysine-specific cysteine proteases and activate apoptosis-like cell death in yeast. **The Journal of Biological Chemistry**, v. 280, n. 15, p. 14691-14699, 2005.

WHISSTOCK, J.C.; PIKE, R. N.; JIN, L.; SKINNER, R.; PEI, X.Y.; CARRELL, R. W.; LESK, A. M. Conformational Changes in Serpins: II. The Mechanism of Activation of Antithrombin by Heparin. **Journal of Molecular Biology**, v. 301, n. 5, p.1287–1305, 2000.

WHISSTOCK, J. C.; BOTTOMLEY, S. P. Molecular gymnastics: serpin structure, folding and misfolding. **Current Opinion in Structural Biology**, v. 16, n. 6, p. 761–768, 2006.

WHISSTOCK, J. C.; BIRD, P. I. An essential role of maspin in embryogenesis and tumor suppression-letter. **Cancer research**, v. 77, n. 18, p. 5207, 2017.

WILLIAMS, B.; DICKMAN, M. Plant programmed cell death: can't live with it; can't live without it. **Molecular Plant Pathology**, v. 9, p.531-44, 2008.

YANG, L.; TADA, Y.; YAMAMOTO, M. P.; ZHAO, H.; YOSHIKAWA, M.; TAKAIWA, F. A transgenic rice seed accumulating an anti-hypertensive peptide reduces the blood pressure of spontaneously hypertensive rats. **FEBS Letters**, v. 580, n. 13, p. 3315- 3320, 2006.

YOO, B. C.; AOKI, K.; XIANG, Y.; CAMPBELL, L. R.; HULL, R. J.; XOCONOSTLE-CAZARES, B.; MONZER, J.; LEE, J. Y.; ULLMAN, D. E; LUCAS, W. J. Characterization of *Cucurbita maxima* phloem serpin-1 (CmPS1) — a developmentally regulated elastase inhibitor. **Journal of Biological Chemistry**, v. 275, n. 45, p. 35122–35128, 2000.

ZHANG, C.; FREDDOLINO, P. L.; ZHANG, Y. COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. **Nucleic Acids Research**, v. 45: p. W291-299, 2017.

ZHANG, H.; FEI, R.; XUE, B.; YU, S.; ZUOMING, Z.; ZHONG, S.; GAO, Y.; ZHOU, X. Pnserpin: A Novel Serine Protease Inhibitor from Extremophile *Pyrobaculum neutrophilum*. **International Journal of Molecular Sciences**, v. 18, n. 1, p. 113, 2017.

ZHANG, Y.; SKOLNICK, J. TM-align: A protein structure alignment algorithm based on TM-score. **Nucleic Acids Research**, v. 33, n. 7, p. 2302-2309, 2005.