

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
DEPARTAMENTO DE GENÉTICA E EVOLUÇÃO

EDUARDO PEREIRA DE SOUZA

**CISTATINAS DE MAQUI: PRODUÇÃO
RECOMBINANTE E INIBIÇÃO DE CISTEÍNO
PEPTIDASES**

SÃO CARLOS - SP

2022

EDUARDO PEREIRA DE SOUZA

CISTATINAS DE MAQUI: PRODUÇÃO RECOMBINANTE E INIBIÇÃO DE CISTEÍNO
PEPTIDASES

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos para obtenção do título de Mestre em Genética Evolutiva e Biologia Molecular.

Orientador: Prof. Dr. Flávio Henrique da Silva

São Carlos - SP

2022



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Eduardo Pereira de Souza, realizada em 15/02/2022.

Comissão Julgadora:

Prof. Dr. Flavio Henrique da Silva (UFSCar)

Profa. Dra. Andrea Soares da Costa Fuentes (UFSCar)

Prof. Dr. Rogério Margis (UFRGS)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

Dedico esse trabalho à memória dos meus avós,
Antônio e Maria Lúcia. Também dedico
aos meus pais, meu irmão e meus avós maternos.
Sem eles, chegar a essa etapa seria impossível.

AGRADECIMENTO

Aos meus pais, Pedro e Erika, e ao meu irmão, Pedro Gustavo, por todo apoio, confiança e incentivo. Aos meus avós, por todo carinho e auxílio, no decorrer da minha vida.

Ao meu orientador Prof. Dr. Flávio Henrique Silva, pela disposição em ajudar, disponibilizar as melhores oportunidades, esclarecer minhas dúvidas. Agradeço por acreditar no meu potencial, pela confiança e paciência.

À Professora Dra. Andrea Soares da Costa Fuentes pelo apoio e dicas nos ensaios fluorimétricos

Ao Professor Dr. Ariel D. Arencibia Rodriguez, pela colaboração nesse trabalho e pela disponibilização do material biológico para que ele fosse realizado.

A Dra. Milene Ferro pela realização das análises de bioinformática.

Aos companheiros de trabalho do Laboratório de Biologia Molecular. À Priscila por todos os ensinamentos e conselhos, muito importantes para a minha trajetória científica. À Thais, pela ajuda em diversas etapas desse trabalho e pela companhia. A Iara, Yan, Bianca, Josilene, Gabrieli, Felipe e Karen, pelas conversas, risadas e auxílios diversos.

Aos amigos de Botucatu, que apesar da distância, sempre me incentivaram e me animaram.

À Universidade Federal de São Carlos (UFSCar) e ao Programa de Pós-graduação em Genética Evolutiva e Biologia Molecular (PPGGEv) pelo suporte e infraestrutura oferecida para a realização deste trabalho.

Ao CNPq, pela bolsa de pesquisa.

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

E a todos que contribuíram para essa pesquisa direta ou indiretamente de alguma forma.

RESUMO

Cistatinas são inibidores competitivos, do tipo *tight-binding*, de cisteíno peptidases. Em plantas, elas atuam regulando diversos mecanismos fisiológicos, bem como protegendo a planta de estresses bióticos e abióticos, o que as torna de uso potencial para aplicações biotecnológicas na agricultura. Em humanos, as cistatinas estão associadas ao controle das cisteíno-catepsinas, que, quando desreguladas, podem desencadear patologias diversas. Desse modo, a intervenção com cistatinas exógenas, como as fitocistatinas, vêm sendo estudada como possível terapia contra diversos tipos de cânceres ou doenças inflamatórias, por exemplo. Diante da grande gama de aplicações das cistatinas, é de relevância a busca de novas variantes em espécies vegetais diversas. O maqui (*Aristotelia chilensis*) é uma planta nativa da América do Sul, intimamente associada à cultura da população chilena. Atualmente, ela vem ganhando destaque devido ao seu alto conteúdo de antocianinas, sendo considerada uma superfruta. Como é uma espécie pouco estudada, pouco se sabe sobre o potencial biotecnológico das suas proteínas, até então, pouco investigadas. Dessa forma, este trabalho consistiu na identificação, produção recombinante e caracterização das cistatinas de maqui. A partir de um transcriptoma da planta cultivada em biorreatores de imersão temporária, foram identificadas seis sequências, das quais cinco foram obtidas por amplificação do cDNA. Estas foram clonadas em vetor de expressão pET-28a para a produção recombinante em *Escherichia coli* Rosetta (DE3). As proteínas purificadas foram submetidas a ensaios fluorimétricos de inibição de cisteíno peptidases. Todas as cistatinas foram capazes de inibir a papaína, com K_i de 7,13, 1,42, 3,29, 2,99, 5,05 nM para as MaquiCPIs de 1 a 5 respectivamente, indicando que foram expressas em conformação correta. Elas também foram capazes de inibir a Catepsina L humana de maneira eficiente (K_i de 0,34, 0,33, 0,38, 0,57 e 1,25 nM), mas somente as MaquiCPIs 1, 2 e 3 foram capazes de inibir a Catepsina B humana na ordem de nanomolar (K_i de 35,74, 20,97 e 21,94 nM, respectivamente), as demais inibiram a proteína na ordem de micromolar. A capacidade das maquicistatinas em inibir as catepsinas humanas refletem em um potencial no controle de diversas patologias.

Palavras-chave: *Aristotelia chilensis*. Catepsina. Fitocistatinas. Inibição *tight binding*. Maquicistatina.

ABSTRACT

Cystatins are tight binding competitive inhibitors of cysteine peptidases. In plants, they act regulating different physiological mechanisms, as well as defending from biotic and abiotic stresses, making them of potential use for biotechnological applications in agriculture. In humans, cystatins are associated with the control of cysteine-cathepsins, which, when deregulated, can trigger distinct pathologies. Then, the intervention with exogen cystatins, like the phytocystatins, have been studied as a therapy for cancer or inflammatory diseases, for example. In the face of the wide variety of cystatins applications, the search for new variants in different plant species is relevant. The maqui berry (*Aristotelia chilensis*) is a native plant from South America, closely related to the Chilean culture. Recently, it has been coming into light due to its high anthocyanin content, and now the species is considered a superfruit. As an understudied species, little is known about the biotechnological potential of its proteins, until then, underinvestigated. In light of this, this work consisted in the identification, recombinant production and characterization of maquicystatins. From a transcriptome of the plant cultivated in Temporary Immersion Bioreactor, six sequences were identified, from which five were obtained by cDNA amplification. These were cloned in pET-28a expression vector for production in *Escherichia coli* Rosetta (DE3). The purified proteins were submitted to fluorimetric assays of inhibition against cysteine peptidases. All cystatins were able to inhibit papain with K_i of 7.13, 1.42, 3.29, 2.99, 5.05 nM, for MaquiCPIs 1 to 5, respectively, proving that they were expressed in the right conformation. They were also able to inhibit human cathepsin L efficiently (K_i of 0.34, 0.33, 0.38, 0.57 and 1.25 nM), but only MaquiCPIs 1, 2 and 3 could inhibit human cathepsin B (K_i of 35.74, 20.97 e 21.94 nM, respectively) in nanomolar order, while the others inhibited the protein in micromolar order. The capacity of maquicystatins in inhibiting human cathepsin indicates a potential in controlling distinct pathologies.

Keywords: *Aristotelia chilensis*. Cathepsin. Maquicystatin. Phytocystatins. Tight binding inhibition.

LISTA DE FIGURAS

Figura 1: Modelo de Schechter e Berger (1970) para os subsítios das peptidases.	16
Figura 2: Estruturas tridimensionais da catepsina L, catepsina B e papaína.	18
Figura 3: Modelo estrutural da orizacistatina I (PDB:1EQK).	21
Figura 4: Determinação da atividade enzimática (A) e da constante de inibição (Ki) (B)	35
Figura 5: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-1.	38
Figura 6: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-2.	39
Figura 7: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-3.	40
Figura 8: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-4.	41
Figura 9: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-5.	42
Figura 10: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-6.	43
Figura 11: Predição de prováveis peptídeos sinais nas maquicistatinas.	46
Figura 12: Predição da estrutura secundária das maquicistatinas.	49
Figura 13: Alinhamento entre MaquiCPI-2 e MaquiCPI-3.	50
Figura 14: Alinhamento das sequências de aminoácidos das MaquiCPIs pelo software Muscle.	52
Figura 15: Alinhamento entre a MaquiCPI-1 e cistatinas de outras espécies vegetais.	54
Figura 16: Alinhamento entre a MaquiCPI-2 e cistatinas de outras espécies vegetais.	54
Figura 17: Alinhamento entre a MaquiCPI-3 e cistatinas de outras espécies vegetais.	54
Figura 18: Alinhamento entre a MaquiCPI-4 e cistatinas de outras espécies vegetais.	55
Figura 19: Alinhamento entre a MaquiCPI-5 e cistatinas de outras espécies vegetais.	55
Figura 20: Alinhamento entre a MaquiCPI-6 e cistatinas de outras espécies vegetais.	55
Figura 21: Árvore filogenética das maquicistatinas.	56
Figura 22: Árvore filogenética das cistatinas de maqui e de arroz (<i>Oriza sativa</i> spp japonica).	57
Figura 23: Amplificação das ORFs das maquicistatinas.	58
Figura 24: Confirmação da clonagem pela clivagem do plasmídeo pET28a+inserto.	59
Figura 25: Expressão recombinante e purificação das maquicistatinas.	61
Figura 26: Representação da estrutura em dímero, por domain swap, da canecystatin-1.	65
Figura 27: Esquema simplificado da transição entre o monômero fechado para o dímero por domain swap.	65
Figura 28: Inibição da Papaína (<i>Carica papaya</i>) pelas cistatinas recombinantes.	66

Figura 29: Inibição da Catepsina L humana pelas cistatinas recombinantes.	67
Figura 30: Inibição da Catepsina B humana pelas cistatinas recombinantes.	68
Figura 31: Clonagem e expressão da PLPro de SARSCoV-2.	90

LISTA DE TABELAS

Tabela 1: Sequências dos primers para amplificar as sequências codificantes das cistatinas de maqui.	30
Tabela 2: Temperaturas de hibridização utilizadas nas ampliações de cada uma das sequências codificantes das cistatinas de maqui.	31
Tabela 3: Tamanho das ORFs e massas moleculares teóricas das cistatinas de maqui.	44
Tabela 4: Características das MaquiCPIs com ^(a) e sem ^(b) sequência sinalizadora.	50
Tabela 5: Similaridade entre as MaquiCPIs obtidas por BLASTP.	51
Tabela 6: Identidade e similaridade da estrutura primária entre as maquicistatinas e outras espécies vegetais.	53
Tabela 7: Características das MaquiCPIs com ^(a) e sem ^(b) a his-tag N-terminal.	60
Tabela 8: Rendimento da produção das diferentes maquicistatinas.	62
Tabela 9: Rendimento por litro de cultura de cistatinas de diferentes espécies	64
Tabela 10: Constantes de inibição (K _i) das maquicistatinas contra a papaína e as catepsinas B e L humanas.	69

LISTA DE ABREVIATURAS E SIGLAS

Asn (N)	asparagine
ATP	trifosfato de adenosina
BLAST	Basic Local Alignment Search Tool
BSA	albumina do soro bovino, do inglês <i>bovine serum albumin</i>
CaCl ₂	cloreto de cálcio
cDNA	DNA complementar
CPI	inibidor de cisteína peptidase, do inglês <i>cysteine peptidase inhibitor</i>
Cys (C)	cisteína
DNA	ácido desoxirribonucleico
dNTP	desoxirribonucleotídeos trifosfatos
D.O.	densidade ótica
DTT	ditiotreitól
EC	Enzyme Commission
g	gramas
Gln (Q)	glutamina
Gly (G)	glicina
h	horas
HCl	ácido clorídrico
His (H)	histidina
IPTG	isopropil β-D-1-tiogalactopiranosida
KCl	cloreto de potássio
kDa	quilodalton
kg	quilograma
KH ₂ PO ₄	fosfato de potássio monobásico
K _{iapp}	constante de inibição aparente
K _i	constante de inibição
K _M	constante de Michaelis
L	litro
LB	caldo de lisogenia, do inglês <i>lysogeny broth</i>

M	molar
mg	miligrama
MgCl ₂	cloreto de magnésio
min	minuto
mL	mililitro
mM	milimolar
Na ₂ HPO ₄	fosfato de sódio dibásico
NaCl	cloreto de sódio
ng	nanograma
ORF	fase aberta de leitura, do inglês <i>open reading frame</i>
pb	pares de base
PBS	tampão fosfato salino
PCR	reação em cadeia da polimerase, do inglês <i>polymerase chain reaction</i>
pH	potencial hidrogeniônico
pI	ponto isoelétrico
RNA	ácido ribonucleico
rpm	rotações por minuto
s	segundos
SDS-PAGE	eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio
TIB	biorreator de imersão temporária, do inglês <i>temporary imersion bioreactor</i>
U	unidade enzimática
X-gal	5-bromo-4-cloro-3-indoxil-β-D-galactopiranosídeo
Z-FR-MCA	carbобенzoxi-Phe-Arg-7-amino-4- metil-coumarina
° C	graus Celsius
μg	micrograma
μL	microlitro
μM	micromolar

SUMÁRIO

1	INTRODUÇÃO	15
1.1	PEPTIDASES	15
1.1.1	Cisteíno peptidases	16
1.1.2	Catepsinas humanas	17
1.1.1.1	Catepsina L	18
1.1.1.2	Catepsina B	19
1.2	CISTATINAS	20
1.2.1	Fitocistatinas	21
1.3	<i>Aristotelia Chilensis</i>	24
2	OBJETIVOS	27
2.1	OBJETIVO GERAL	27
2.2	OBJETIVOS ESPECÍFICOS	27
3	METODOLOGIA	28
3.1	IDENTIFICAÇÃO E CARACTERIZAÇÃO DAS SEQUÊNCIAS	28
3.1.1	Identificação das fitocistatinas	28
3.1.2	Análise <i>in silico</i> das sequências	29
3.1.3	Análise filogenética	29
3.1.3.1	Filogenia entre as maquicistatinas	29
3.1.3.2	Filogenia entre as cistatinas de maqui e de <i>Oriza sativa</i>	30
3.2	AMPLIFICAÇÃO E CLONAGEM DAS ORFS	30
3.2.1	Isolamento das ORFs que codificam as cistatinas	30
3.2.2	Construção de vetores de expressão	31
3.3	EXPRESSÃO DAS CISTATINAS RECOMBINANTES	33
3.3.1	Expressão das cistatinas recombinantes em células de <i>E. coli</i>	33
3.3.2	Lise bacteriana e teste de solubilidade	33

3.3.3	Purificação das cistatinas recombinantes	34
3.4	INIBIÇÃO DA ATIVIDADE DE CISTEÍNO PEPTIDASES UTILIZANDO MAQUICISTATINAS	34
4	RESULTADOS E DISCUSSÃO	37
4.1	IDENTIFICAÇÃO DAS SEQUÊNCIAS E ANÁLISE <i>IN SILICO</i>	37
4.2	ANÁLISE FILOGENÉTICA	56
4.3	ISOLAMENTO DAS ORFS E CONSTRUÇÃO DE VETORES DE EXPRESSÃO	57
4.4	EXPRESSÃO E PURIFICAÇÃO DAS CISTATINAS RECOMBINANTES	59
4.5	ENSAIOS DE INIBIÇÃO DE CISTEÍNO PEPTIDASES	65
5	CONCLUSÕES	71
6	PERSPECTIVAS E TRABALHOS SENDO REALIZADOS	72
	REFERÊNCIAS	73
	APÊNDICE A – Produção da <i>Papain like protease</i> de SARSCoV-2	89
	APÊNDICE B – Sequências das MaquiCPIs	92
	ANEXO 1 – Mapa do vetor pTZ57R_T	98
	ANEXO 2 – Mapa do vetor pET28a	99

1 INTRODUÇÃO

1.1 PEPTIDASES

Peptidases, proteases ou enzimas proteolíticas, são proteínas que catalisam a hidrólise de ligações peptídicas (BARRETT; MCDONALD, 1986), sendo assim categorizadas pelo número EC 3.4 (hidrolases que atuam sobre ligações peptídicas). Elas são essenciais para a existência de todos os tipos de organismos vivos. Estão associadas às mais básicas funções, como a quebra de proteínas para reciclo de aminoácidos e produção de novas proteínas, além de funções mais refinadas como o processamento de um polipeptídeo para a sua ativação. Assim, estão relacionadas a diversos processos bioquímicos e fisiológicos (BARRETT, 2000).

As enzimas proteolíticas podem ser divididas em dois grupos em função do sítio de clivagem. Nesse sentido, podem ser exopeptidases (EC 3.4.11-19), que liberam um ou poucos aminoácidos da região N ou C-terminal, ou endopeptidases (EC 3.4.21-25 e EC 3.4.99) que clivam internamente (BARRETT, 2001; BARRETT; RAWLINGS; WOESSNER, 2012).

As peptidases podem ser classificadas de acordo com o mecanismo de catálise, sendo assim agrupadas em seis grupos. As serino, cisteíno e treonino peptidases apresentam, em seu sítio catalítico, aminoácidos altamente nucleofílicos que dão os nomes das suas classes. Seu mecanismo catalítico é denominado catálise covalente, ou seja, é formado um intermediário covalente entre a enzima e fragmento do seu substrato. Já na catálise pelas aspartil, metalo e glutamil peptidases, o responsável pelo ataque nucleofílico não é um resíduo de aminoácido da enzima, pois quem o realiza é uma molécula ativada de água. Dessa forma, não é observado um intermediário covalente, caracterizando uma catálise ácido-básica. Nesses casos, os resíduos de aspartato e glutamato, ou um íon metálico, atuam como bases (BARRETT; RAWLINGS; WOESSNER, 2012; RAWLINGS *et al.*, 2018).

O sítio ativo de uma peptidase, de acordo com as estruturas cristalográficas, é comumente localizado em uma cavidade na superfície da macromolécula entre domínios estruturais adjacentes. A especificidade do substrato está sujeita às propriedades dos sítios de ligação arranjados ao redor da cavidade catalítica responsável pela quebra da ligação peptídica (BARRETT, 2001). Nesse sentido, foi proposto um modelo (Figura 1) que é utilizado para descrever a especificidade de qualquer uma das classes de peptidases (BERGER; SCHECHTER, 1970). O esquema mostra como os locais de ligação de substrato são distribuídos ao redor do sítio catalítico. Os sítios em direção ao N-terminal são denominados

S1, S2, Sn, enquanto aqueles em direção ao C-terminal são apresentados como S1', S2', Sn'. Os aminoácidos, por sua vez, recebem a mesma numeração, mas são representados pela letra P. Assim, P1 e P1' são os aminoácidos que serão separados ao se romper a ligação peptídica.

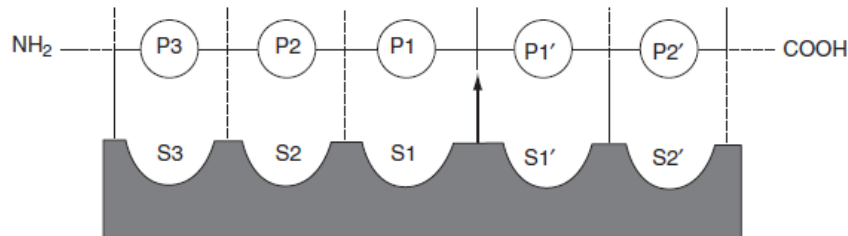


Figura 1: Modelo de Schechter e Berger (1970) para os subsítios das peptidases. Os sítios em que os aminoácidos da proteína alvo interagem são representados por S_n e os aminoácidos são representados por P_n . A seta ao centro indica a ligação peptídica a ser clivada.

1.1.1 Cisteíno peptidases

Distribuídas filogeneticamente pelos mais diversos organismos, as cisteíno ou tiol peptidases correspondem às enzimas proteolíticas que possuem um resíduo de cisteína em seu sítio ativo (RAWLINGS; BARRETT, 1994).

As peptidases cisteínicas são agrupadas em dez clãs (ou superfamílias), de acordo com estrutura e relações evolutivas, além da sequência e a natureza dos aminoácidos do sítio ativo. O clã CA, por exemplo, possui uma tríade catalítica formada por cisteína, histidina e aspartato (ou asparagina), nessa sequência. Ele corresponde ao conjunto de cisteíno peptidases mais estudado e seus representantes apresentam estrutura similar a papaína. Esse apresenta 44 famílias, incluindo a bem elucidada C1, ou família da papaína, que é subdividida nas subfamílias C1A e C1B. A primeira agrupa peptidases que contém pontes dissulfeto e que se acumulam em vesículas, em vacúolos ou no apoplasto, enquanto as representantes da segunda são localizadas no citoplasma e não possuem essas ligações entre resíduos de cisteína. Em plantas há apenas exemplares C1A (RAWLINGS *et al.*, 2018).

Essas proteínas são sintetizadas na forma de pré-propeptidases. Para a sua ativação, o pró-domínio auto inibitório dobra-se para trás na fenda do sítio catalítico e é removido (TAYLOR *et al.*, 1995). A estrutura terciária madura dos exemplares da família C1 é bilobada, apresentando um domínio L (*left*, da esquerda), que corresponde a três α -hélices da região N-terminal e um domínio R (*right*, da direita) que possui um motivo barril- β na porção C-terminal. Na maioria das enzimas deste grupo, a estrutura é estabilizada por três ligações dissulfeto, uma

no domínio R e duas no domínio L. Entre esses dois lobos, encontra-se o sítio catalítico, com a cisteína catalítica à esquerda e a histidina à direita (DRENTH *et al.*, 1968; MCGRATH, 1999).

O mecanismo de hidrólise por cisteíno peptidases envolve diversos passos. A primeira etapa corresponde à associação da enzima livre com o substrato formando o complexo de Michaelis. Em sequência, ocorre a formação do intermediário covalente acil-enzima, resultante de um ataque nucleofílico pelo grupo tiol do sítio ativo ao carbono carbonila da ligação amida. Assim, ocorre a saída do primeiro produto da reação (R-NH₂). Em seguida, o complexo acil-enzima reage com uma molécula de água para formar o segundo produto (R-COOH), na etapa de deacilação. Vários intermediários e/ou estados de transição são preditos durante essa rota. No entanto, a exata natureza do mecanismo de reação vem sendo debatido (STORER; MÉNARD, 1994, 2013).

1.1.2 Catepsinas humanas

O termo catepsina, o qual é derivado do grego *kathepsin* (digerir), foi proposto para peptidases que possuíssem atividade em ambientes ligeiramente ácidos (WILLSTÄTTER; BAMANN, 1929). Subsequentemente, o termo foi designado às serino-peptidases catepsinas A e G, às aspártico-peptidases catepsinas D e E, e às cisteíno-catepsinas lisossomais. Em humanos, esse último grupo se apresenta na forma de 11 peptidases distintas: as catepsinas B, C, F, H, K, L, O, S, V, X, W (ROSSI *et al.*, 2004).

Inicialmente, as cisteíno-catepsinas eram consideradas enzimas intracelulares responsáveis pela atividade proteolítica não específica em massa no ambiente ácido e redutor dos compartimentos lisossomais e endossomais, onde degradam proteínas intra e extracelulares (TURK; TURK; TURK, 2000; TURK, 2001). No entanto, essa visão vem sendo alterada, pois é notável sua atuação em ambiente extracelular, indicando que o pH não seria o único responsável por sua atividade (TURK *et al.*, 2012; VIDAK *et al.*, 2019).

Em função das características conservadas das suas estruturas tridimensionais, as peptidases da família da papaína nos mamíferos são divididas em duas subfamílias diferentes, designadas como do tipo catepsina B (*cathepsin B-like*) e do tipo catepsina L (*cathepsin L-like*). Elas diferem principalmente em relação a presença de um segmento adicional na estrutura das catepsinas *B-like* e quanto à especificidade do substrato, o que está associado ao subsítio S2 na família da papaína (Figura 2) (COULOMBE *et al.*, 1996; FUJISHIMA *et al.*, 1997).

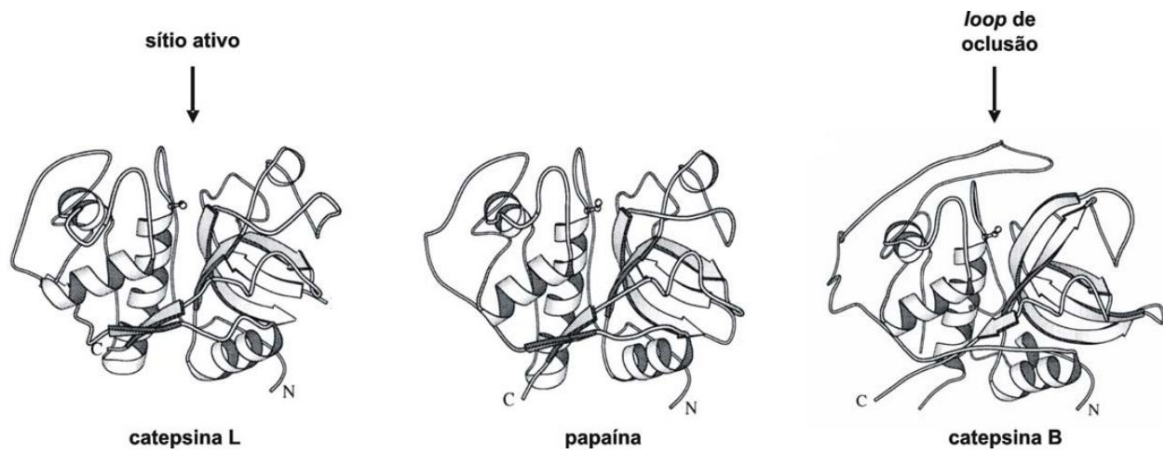


Figura 2: Estruturas tridimensionais da cathepsina L, cathepsina B e papaína. A figura mostra a presença de um segmento peptídico adicional na estrutura da cathepsina B, configurando um *loop* de oclusão que a diferencia das outras cisteíno peptidases da família da papaína. N e C correspondem às extremidades N-terminal e C-terminal das proteínas. O resíduo de cisteína está representada pelo modelo *ball-stick*. Adaptado de Fujishima *et al.*, 1997.

1.1.1.1 Cathepsina L

O gene da cathepsina L humana CTSL (EC:3.4.22.15; Uniprot: P07711) está localizado na posição 9q21-q22 do cromossomo 9, codificando uma sequência peptídica de 333 aminoácidos (FAN *et al.*, 1989). A região codante inclui um peptídeo sinal N-terminal, dois propeptídeos e dois peptídeos maduros, contendo uma cadeia leve e uma pesada. Após a transcrição, o peptídeo sinal é removido de maneira co-traducional no polissomo resultando na pró-cathepsina L, que é translocada ao retículo endoplasmático onde passa por N-glicosilação com açúcares ricos em manose. Estes são fosforilados em *cis* no complexo de Golgi (LANG *et al.*, 1984). A partir da rota endolisossomal a pró-cathepsina L é destinada ao lisossomo, onde os resíduos de fosfato são removidos (HASHIMOTO; KONDO; KATUNUMA, 2015). Nesse ambiente, a ativação em cathepsina L madura ocorre pela remoção do propeptídeo tanto por autocatálise (JERALA *et al.*, 1998) ou pela ação da cathepsina D, uma aspartil peptidase (NISHIMURA *et al.*, 1989). Isso resulta na cathepsina L ativa de duas cadeias, leve e pesada, conectadas por ligações dissulfeto (DANA; PATHAK, 2020).

A cathepsina L atua principalmente como endopeptidase, realizando clivagens preferencialmente associadas a ligações peptídicas de substratos com resíduos aromáticos na região P2 (FUJISHIMA *et al.*, 1997). A estrutura tridimensional da pró-cathepsina L demonstrou que seu propeptídeo estabelece interações hidrofóbicas com a região correspondente à fenda do sítio ativo da enzima madura. Nesse caso, os subsítios S1 e S3, são ocupados pelo propeptídeo, porém em direção oposta àquela apropriada para clivagem (COULOMBE *et al.*, 1996).

Diferentes formas da catepsina L foram isoladas dos mais diversos tipos celulares em seus ambientes intra e extracelulares. Em localizações celulares específicas, elas atuam em papéis fundamentais em uma grande variedade de atividades funcionais. Nota-se, por exemplo, que, além da degradação de proteínas celulares na rota endolisossomal, essa peptidase também participa de processos de autofagia (DANA; PATHAK, 2020; REISER; ADAIR; REINHECKEL, 2010). Diante da sua ubiquidade no corpo humano, possui papel chave em diversos sistemas fisiológicos, podendo estar associada, em situações anormais, ao surgimento de determinadas patologias. Dentre elas, se incluem fibrose no fígado, diabetes do tipo I e II, doenças imunes, cardíacas, ósseas e renais, além de diversos tipos de cânceres (DANA; PATHAK, 2020).

1.1.1.2 Catepsina B

O gene da catepsina B humana CTSB (EC:3.4.22.1; Uniprot: P07858) está localizado na posição 8p22 do cromossomo 8, codificando uma sequência peptídica de 339 aminoácidos (WANG *et al.*, 1987). Além do conteúdo referente à enzima madura, o polipeptídeo é sintetizado com peptídeo sinal, pró-região e uma extensão C-terminal de 6 resíduos (CHAN *et al.*, 1986; WIEDERANDERS, 2003).

A catepsina B possui preferência por resíduos hidrofóbicos na posição P2 do substrato, no entanto, é a única que comporta um resíduo de arginina nessa posição, por possuir um resíduo de ácido glutâmico (Glu245) no subsítio S2 (HASNAIN *et al.*, 1993). Essa enzima, além de poder atuar como endopeptidase, clivando ligações peptídicas internas, pode agir como exopeptidase (atividade carboxipeptidase) (KEPPLER; SLOANE, 1996; MUSIL *et al.*, 1991). Essa atividade dupla é decorrente da presença de um loop de oclusão (Figura 2) que interfere no acesso de substratos ao sítio ativo (ILLY *et al.*, 1997). Quando em pH ácido, o *loop* de oclusão bloqueia parcialmente o sítio ativo da molécula prevenindo a entrada de substratos maiores, mas permitindo a entrada de substratos sintéticos ou a região carboxiterminal das proteínas. Enquanto, em pH ácido, a catepsina B possui atividade carboxipeptidase, em pH neutro, ela tem seu *loop* deslocado, de modo que não bloqueia o sítio ativo e permite a entrada de substratos maiores, desempenhando atividade endopeptidase (ILLY *et al.*, 1997; QURAIISHI *et al.*, 1999).

A existência desse *loop* de oclusão na peptidase madura é incompatível com a inibição por membros da família das cistatinas, inibidores proteicos reversíveis de cisteíno peptidases, pois eles se ligam aos sítios S e S' da enzima (ILLY *et al.*, 1997). Essa inibição só se torna

possível devido a um mecanismo de dois passos, que está associado ao deslocamento do *loop* de oclusão antes da ligação à fenda do sítio ativo (NYCANDER *et al.*, 1998; PAVLOVA *et al.*, 2000). Diante desse fato, a maioria dos inibidores naturais de cisteíno peptidases são menos efetivos contra a catepsina B.

Sob condições fisiológicas normais, a catepsina B está de maneira primária associada ao ciclo de proteínas intra e extracelulares, mantendo a homeostase metabólica dentro das células. Outras funções celulares incluem a regulação da ativação de pró-hormônios e pró-enzimas, processamento de antígenos, resposta inflamatória, remodelação tecidual e apoptose (MORT; BUTTLE, 1997; REISER; ADAIR; REINHECKEL, 2010). Diante disso, variações nos níveis de expressão e na atividade da catepsina B estão correlacionadas a uma gama de patologias incluindo câncer, artrite, pancreatite, doenças cardiovasculares e doença de Alzheimer (CAVALLO-MEDVED; MOIN; SLOANE, 2011).

1.2 CISTATINAS

Cistatinas são inibidores proteicos reversíveis de cisteíno peptidases semelhantes a papaína (*papain-like*, subfamília C1A no MEROPS), que apresentam mecanismo competitivo, ou seja, interagem com o sítio ativo na fenda da peptidase alvo (BARRETT, 1987; BARRETT *et al.*, 1986; RAWLINGS *et al.*, 2018). O seu alto poder inibitório é explicado pela interação do tipo *tight-binding*, que ocorre devido a três pontos de contato entre a cistatina e a peptidase. Dois desses pontos correspondem a *loops* em *hairpin* formados entre fitas beta antiparalelas. O primeiro *loop* de ligação contém o motivo QxVxG, que é altamente conservado, enquanto o último apresenta um resíduo de triptofano. O terceiro ponto de contato é um resíduo de glicina conservado na extremidade N-terminal, que se acomoda nos subsítios S1 e S2 da enzima, mas em conformação que a torna invulnerável ao ataque proteolítico (ABRAHAMSON *et al.*, 1987; RAWLINGS; BARRETT, 1990; TURK; BODE, 1991).

A estrutura terciária das cistatinas, de modo geral, é caracterizada por apresentar uma alfa-hélice central de cinco voltas, ao redor da qual, cinco folhas beta antiparalelas estão torcidas e enroladas, além de um elemento alfa-helicoidal parcial (Figura 3) (BENCHABANE *et al.*, 2010; TURK; BODE, 1991).

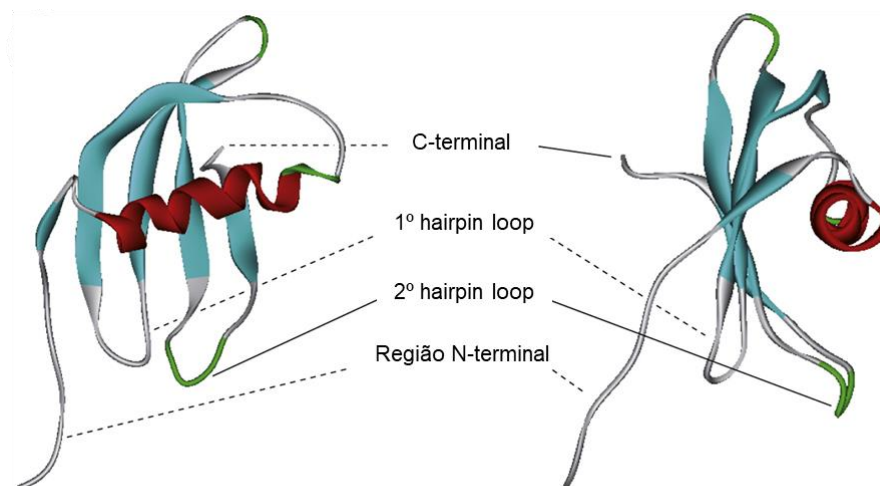


Figura 3: Modelo estrutural da orizacistatina I (PDB:1EQK). Adaptado de Benchabane *et al.*, 2010. A figura mostra a folha beta com 5 fitas em azul e a α -hélice com o motivo conservado LARFAVDEHN em vermelho. Os motivos essenciais associados a inibição estão indicados: na região N-terminal se encontram os resíduos conservados de glicina seguido pelo primeiro (QxVxG) e segundo loop (PW).

A superfamília das cistatinas foi subdividida, a priori, em três famílias baseadas na identidade/similaridade das sequências, na presença e posição de ligações dissulfeto e na massa molecular (BARRETT, 1987; TURK; BODE, 1991). A família das estefinas consiste em proteínas sem ligações dissulfeto ou carboidratos, com massa aproximada de 11 kDa. A segunda família, denominada simplesmente “cistatinas”, agrupa proteínas com cerca de 115 aminoácidos e quatro resíduos de cisteína formando duas ligações dissulfeto, e que podem ser glicosiladas ou fosforiladas. A família dos cininogênios, por sua vez, é composta por grandes glicoproteínas de cerca de 60-120 kDa, que contém domínios homólogos à família 2 das cistatinas (BARRETT *et al.*, 1986; TURK; BODE, 1991).

Formando um clado à parte, as cistatinas de plantas divergem das de animais, por apresentarem peculiaridades estruturais e genômicas, bem como diversidade intrínseca. Por essa razão, posteriormente, foi criada a família das fitocistatinas (MARGIS; REIS; VILLERET, 1998).

1.2.1 Fitocistatinas

Quando comparadas às cistatinas animais, além de apresentarem os motivos funcionais clássicos, as fitocistatinas apresentam um motivo adicional conservado na alfa-hélice, de função desconhecida, que possui a sequência [LVI]-[AGT]-[RKE]-[FY]-[AS]-[VI]-X-[EDQV]-[HYFQ]-N (MARGIS; REIS; VILLERET, 1998).

As fitocistatinas podem ser classificadas em grupos de acordo com seus aspectos estruturais e funcionais, promovidos ao longo do tempo por meio de mutações e duplicações gênicas de cistatinas de domínio único, que acarretam divergências funcionais (BENCHABANE *et al.*, 2010). Este fenômeno permitiu o surgimento das multicistatinas, que são polipeptídeos compostos por diversos domínios de cistatinas *in tandem*, os quais sofreram o processo de evolução adaptativa, gerando hipervariabilidade intramolecular e diversidade funcional contra peptidases do tipo papaína (GOULET *et al.*, 2008; KIGGUNDU *et al.*, 2006). O processo de duplicação gênica também foi responsável pelo surgimento de cistatinas de cerca de 23 kDa com C-terminal estendido (MARTÍNEZ *et al.*, 2005). O domínio C-terminal dessas proteínas teria sido submetido a um remodelamento funcional e estrutural significativo, desencadeando a perda da atividade contra peptidases do tipo papaína e a aquisição de outras funções como a inibição de cisteíno peptidases do tipo legumaína (Família C13). Esse processo está associado a um motivo de sequência SNSL, o que também foi observado em cistatinas animais como a cistatina M/E humana (ALVAREZ-FERNANDEZ *et al.*, 1999; MARTINEZ *et al.*, 2007).

A ubiquidade das cistatinas no reino vegetal, a habilidade da maioria delas em inibir cisteíno peptidases, além da alta conservação da estrutura e dos motivos inibitórios, indicam que elas desempenham papéis vitais nesses organismos (BENCHABANE *et al.*, 2010).

De maneira endógena, a atuação das cistatinas é bastante relevante para a fisiologia e desenvolvimento do vegetal, devido à atividade de suas cisteíno peptidases que necessitam ser reguladas em diversos processos celulares. Em sementes, no início do desenvolvimento ou em órgãos de armazenamento vegetativo, as cistatinas são sintetizadas ativamente para superar o número de cisteíno peptidases e promover o acúmulo de proteínas. Assim, uma alta relação cistatina/Cys-peptidases em tecidos em dormência, permite que o pool de proteínas armazenadas seja mantido, estando disponível para o crescimento das plântulas durante a germinação. A indução da expressão de genes codificadores de peptidases e a concomitante repressão dos codificadores de cistatinas, após a embebição ou semeadura, causam uma acentuada redução da razão cistatina/Cys-peptidase, induzindo a hidrólise e a mobilização das proteínas de armazenamento. Essa característica se mantém durante o início da germinação (HWANG *et al.*, 2009; VALDÉS-RODRÍGUEZ *et al.*, 2007). Esse balanço é regulado em nível de transcrição pela ação dos fitormônios giberelina, que inibe a síntese de cistatina e estimula a germinação, e ácido abscísico, que estimula a produção de cistatina e inibe a germinação (HWANG *et al.*, 2009; MARTINEZ *et al.*, 2003).

Essa família de inibidores de peptidases ainda apresenta outros efeitos *in planta*. Ela atua nas folhas e outros órgãos metabolicamente ativos para o controle das cisteíno peptidases envolvidas em processos biológicos como o reciclo de proteínas durante a senescência e a regulação de cascatas metabólicas de proteção durante estresse abiótico (SUGAWARA *et al.*, 2002; ZHANG; LIU; TAKANO, 2008). As fitocistatinas já demonstraram serem capazes de induzir defesa frente à seca (TAN *et al.*, 2017), salinidade (TAN *et al.*, 2017), frio (VAN DER VYVER *et al.*, 2003), entre outros estresses.

As cistatinas ainda apresentam efeitos *ex planta*, o que é evidente pela regulação positiva dos genes codificantes de cistatinas frente a ferimentos mecânicos, ácido jasmônico, elicitores fúngicos ou fitofagia (LIANG *et al.*, 2015; PERNAS; SÁNCHEZ-MONGE; SALCEDO, 2000). Diversos estudos demonstraram o potencial das fitocistatinas no controle de fungos fitopatogênicos, no entanto, o mecanismo de atuação ainda não está bem esclarecido (SHYU *et al.*, 2011; SOARES-COSTA *et al.*, 2002). Quando se tratam de artrópodes fitófagos, os mecanismos estão melhor elucidados. O mais comum deles se baseia na inibição de suas cisteíno peptidases digestivas. Assim, as peptidases, quando secretadas para digerir o tecido do vegetal, não são capazes de clivar ligações peptídicas, resultando na interrupção prejudicial da assimilação proteica, induzindo à redução no crescimento e ao aumento da mortalidade (ARAI; ABE, 2000; MARTINEZ *et al.*, 2016).

Nota-se, portanto, que as fitocistatinas apresentam grande potencial biotecnológico na agricultura. Elas podem ser utilizadas em variantes vegetais transgênicas, visando a indução de resistência, ou também como uma alternativa bastante útil frente aos pesticidas químicos para o controle de pragas e patógenos no campo (BENCHABANE *et al.*, 2010).

As fitocistatinas, além de promissoras na agricultura, apresentam potenciais aplicações na medicina. Esse fator está, em grande parte, associado ao seu potencial inibitório frente às cisteíno catepsinas humanas (SCHNEIDER *et al.*, 2020; SHIBAO *et al.*, 2021), o que as torna capazes de reverter anomalias funcionais ou quantitativas relacionadas a essas peptidases. Ensaio utilizando uma cistatina de cana-de-açúcar (*Saccharum officinarum*) (CaneCPI-4) mostraram a sua capacidade de inibir o desenvolvimento *in vivo* de melanoma por alterações na angiogênese e pela inibição dos processos de invasão, migração e crescimento dependente de ancoragem (OLIVEIRA *et al.*, 2011). Estudos recentes demonstraram um papel relevante das fitocistatinas no processo inflamatório. Foi observado que as proteínas CsinCPI-2 (de laranja) e CaneCPI-5 (de cana-de-açúcar), são capazes de reduzir a síntese de mediadores pró-inflamatórios (FERREIRA *et al.*, 2021; LEGUIZAMÓN *et al.*, 2019). Além disso, já foram

comprovados efeitos pró-osteogênicos da CsinCPI-2 em polpa dentária (DA COSTA FERNANDES *et al.*, 2021; LEGUIZAMÓN *et al.*, 2019).

Outras aplicações na saúde estão associadas à inibição de patógenos, colaborando para o combate a determinadas patologias. Nesse sentido, já foi observada, por exemplo, a inibição de enzimas (falcipainas) de protozoários como o *Plasmodium falciparum*, pela CaneCPI-4 (MELO *et al.*, 2018). Há também relatos do controle da proliferação do patógeno fúngico *Candida spp.* por uma cistatina de *Cassia leiandra* pela destruição da superfície celular (MELO *et al.*, 2019).

Recentemente, a CaneCPI-5 tem sido aplicada na área da odontologia. Foi observado que ela é capaz de se ligar ao esmalte dentário, reduzindo a erosão dentária (PELÁ *et al.*, 2021; SANTIAGO *et al.*, 2017). Estudos posteriores mostraram que a mesma cistatina era também capaz de promover efeitos antibacterianos e anticárie (ARAUJO *et al.*, 2021).

Diante das diversas possíveis aplicações biotecnológicas das cistatinas de plantas, é evidente a relevância da identificação e caracterização de novos exemplares desses inibidores. Assim, o estudo de fitocistatinas oriundas de espécies pouco exploradas torna-se uma fonte promissora de novas proteínas desta classe.

1.3 *Aristotelia Chilensis*

Nativo da América do Sul, o maqui [*Aristotelia chilensis* (Mol.) Stuntz] é uma espécie vegetal arbustiva e perene, que pode chegar a 5 metros de altura. Pertencente à família Elaeocarpaceae, cresce em matas densas e prefere solos ligeiramente ácidos, moderadamente férteis e bem irrigados. Em condições de umidade adequadas, ela cresce rapidamente, possibilitando a colonização de solos abandonados, queimados ou sobre-explorados, protegendo-os do processo erosivo. É uma planta dióica que produz pequenos frutos, de cerca de 6 mm de diâmetro, de coloração roxa ou preta (HOFFMANN *et al.*, 1992).

Esta espécie vegetal está tradicionalmente associada à cultura da população chilena. Por serem refrescantes, ligeiramente ácidos e adstringentes, seus frutos são comestíveis e também utilizados na manufatura de geleias, sucos e bebidas alcólicas. Os frutos também fazem parte da medicina popular, sendo utilizados no combate a diarreia e a disenteria, assim como as folhas que atuam com efeitos anti-inflamatórios, antiespasmódicos e analgésicos (MISLE *et al.*, 2011).

Atualmente, o maqui vem ganhando destaque devido ao seu conteúdo de fitoquímicos. Estudos que visavam elucidar a composição química do maqui indicaram a presença dos alcalóides indol e quinolina (HE *et al.*, 1997), além dos polifenóis, de maior destaque (MIRANDA-ROTTMANN *et al.*, 2002). Os fenóis são um grupo de princípios ativos orgânicos com anéis aromáticos com substituições de grupos hidroxila, comumente agrupados como polifenóis (DILLARD; GERMAN, 2000). Na espécie, há ênfase às antocianinas, polifenóis hidrossolúveis do grupo dos flavonoides, por seu conteúdo (que chega a 31,2 g/kg) e atividade (BRAUCH *et al.*, 2016; CÉSPEDES *et al.*, 2010). Majoritariamente, as antocianinas presentes correspondem a derivados de delphinidinas e cianidinas (BRAUCH, 2016). A sua intensa coloração, resultante da sua composição, coloca-o em destaque como um potencial pigmento natural, com alta demanda para o mercado europeu, devido ao conhecimento acerca dos males associados aos corantes alimentícios artificiais (MISLE *et al.*, 2011).

O alto conteúdo de antocianinas dos frutos foi relatado como sendo o responsável pelas suas altas propriedades antioxidantes, as mais elevadas entre as “berries” ao redor do mundo (MIRANDA-ROTTMANN *et al.*, 2002). Diante desse fato, o consumo do fruto foi associado a diversos benefícios à saúde, o que rendeu a ele o título de super-fruta (BRAUCH, 2016). Dentro desse contexto, as antocianinas de maqui foram descritas por possuírem efeitos cardioprotetores (CÉSPEDES *et al.*, 2008), antidiabéticos (ROJO *et al.*, 2012) e anti-inflamatórios (CÉSPEDES *et al.*, 2010).

Diante desses benefícios, recentemente, estratégias vêm sendo desenvolvidas a fim de se obter os metabólitos secundários dessa espécie. Nesse contexto, o uso de biorreatores de imersão temporária (TIB, *temporary immersion bioreactor*) demonstrou ser uma tecnologia eficaz. Foi observado o estabelecimento de culturas fotomixotróficas da planta nesse sistema, onde foi induzida a produção dos fenilpropanóides pela adição do fitorregulador ABA. Isso permitiu uma superprodução de metabólitos antioxidantes com atividade biológica semelhante às frutas obtidas de plantas crescendo *ex vitro* (TRENTINI *et al.*, 2021).

Muito vem sendo descrito a respeito dos extratos botânicos do maqui e da amplitude química dos seus metabólitos secundários, entretanto, pouco foi reportado sobre os seus genes e proteínas. As proteínas vegetais são essenciais para a manutenção da homeostase do organismo e podem permitir vantagens adaptativas no que se refere à defesa contra patógenos, pragas e estresses abióticos, entre outras propriedades específicas. Diante disso, elas apresentam uma infinidade de aplicações biotecnológicas em prática e potenciais, abrangendo

diversos setores da indústria, da medicina e da agricultura. Nesse cenário, destaca-se a importância das fitocistatinas.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Identificar, produzir e caracterizar cistatinas de maqui.

2.2 OBJETIVOS ESPECÍFICOS

- Identificar sequências candidatas a cistatinas de *A. chilensis* a partir do transcriptoma da planta cultivada em Biorreator de Imersão Temporária (*Temporary Immersion Bioreactor, TIB*).
- Analisar e caracterizar as sequências proteicas *in silico*.
- Amplificar e clonar cada uma das cistatinas no vetor pET-28a.
- Expressar as proteínas em *Escherichia coli* Rosetta (DE3) e purificá-las.
- Caracterizá-las quanto ao seu poder inibitório contra a cisteíno peptidases.

3 METODOLOGIA

3.1 IDENTIFICAÇÃO E CARACTERIZAÇÃO DAS SEQUÊNCIAS

3.1.1 Identificação das fitocistatinas

Plântulas de *Aristotelia chilensis* foram cultivadas em biorreator de imersão temporária em meio MS (MURASHIGE; SKOOG, 1962) na ausência ou presença do fitorregulador ácido abscísico (ABA). Amostras do material vegetal foram coletadas em diferentes tempos de cultura (0, 24 e 48 horas). O cultivo foi conduzido no Laboratório de Cultivo y Tejido Vegetal, da Universidade Católica de Maule (UCM) (Chile), coordenado pelo Professor Doutor Ariel D. Arencibia, colaborador desse trabalho. A planta inteira cultivada foi utilizada para a extração de RNA com o TRIzol, conforme instruções do fabricante.

O RNA total foi submetido a sequenciamento utilizando o TruSeq Small RNA Library Preparation Kit (Illumina) na plataforma Hi-Seq (Illumina). O sequenciamento foi conduzido no Centro de Genômica Funcional da ESALQ/USP. O processamento dos *reads* e a busca das sequências foi realizado pela nossa colaboradora Dra. Milene Ferro, pós-doutoranda da UNESP, Rio Claro. Os *reads* obtidos foram submetidos à análise de qualidade e a um pré-processamento. Os *contigs* foram montados a partir dos *reads* obtidos do RNA da amostra sem ABA coletada no início do experimento. Dentro da plataforma Omicsbox, os *contigs* foram utilizados como alvos para busca de sequências candidatas a cistatinas. Esse processo consistiu na utilização de 14 sequências de cistatinas de diferentes grupos e espécies vegetais como iscas em uma busca por BLASTP local. Foram utilizadas sequências de *Fragaria chiloensis* (AWO72930.1), *Prunus dulcis* (BBN70259.1), *Musa acuminata malaccensis* (XP_009387215.1, XP_009386948.1), *Solanum lycopersicum* (AAF23126.1), *Vitis cinerea var. helleri x Vitis riparia* (ADD51191.1), *Rosa chinensis* (XP_024185425.1), *Malus prunifolia* (AHZ92270.1), *Camellia sinensis* (ACN85344.1), *Cucumis melo var. makuwa* (KAA0040524.1), *Morus notabilis* (EXC30697.1), *Capsicum annuum* (PHT93956.1) e *Fragaria vesca subsp. vesca* (XP_011468407.1, XM_011463703.1). Alguns critérios foram utilizados para a identificação das sequências alvo: existência dos códons de iniciação e de terminação além da presença de três regiões características das cistatinas: glicina-glicina conservados na região N-terminal, QxVxG no *loop* central e prolina-triptofano conservado na região C-terminal (BODE *et al.*, 1988; TURK; BODE, 1991). Diante disso, foram identificadas seis cistatinas para caracterização.

3.1.2 Análise *in silico* das sequências

As sequências identificadas foram traduzidas pelo programa Translate Tool (GASTEIGER, 2003) e algumas de suas características foram investigadas. A presença de peptídeo sinal foi avaliada pelos programas PrediSi (HILLER *et al.*, 2004), SignalP (ALMAGRO ARMENTEROS *et al.*, 2019), Phobius (KÄLL; KROGH; SONNHAMMER, 2004) e iPSORT (BANNAI *et al.*, 2002). As massas moleculares teóricas foram calculadas pelo programa Compute pI/MW (GASTEIGER, 2003). A estrutura secundária de cada uma das proteínas foi determinada pelo programa PSIPRED (MCGUFFIN; BRYSON; JONES, 2000).

A fim de verificar a semelhança das cistatinas de maqui frente às de outras espécies, foi realizado um BLASTP dentro do banco de dados do NCBI. As sequências de aminoácidos das proteínas mais próximas selecionadas foram alinhadas junto à maquicistatina correspondente pelo método Muscle (EDGAR, 2004) na plataforma Mega 11 (TAMURA; STECHER; KUMAR, 2021). Os alinhamentos foram efetuados com os parâmetros padrão e formatados com o Esript (ROBERT; GOUET, 2014). A mesma metodologia, também foi utilizada para a construção de um alinhamento múltiplo entre as sequências de aminoácidos das seis maquicistatinas.

3.1.3 Análise filogenética

3.1.3.1 Filogenia entre as maquicistatinas

A partir do alinhamento múltiplo entre as sequências de aminoácidos das seis cistatinas de maqui descrito na seção anterior, foi realizada uma análise filogenética dentro da plataforma Mega 11 (TAMURA; STECHER; KUMAR, 2021). Os *gaps* presentes no alinhamento foram removidos, de modo que somente as posições em que havia pelo menos 75% de cobertura foram mantidas. Utilizando essa plataforma, a melhor modelo de substituição de aminoácidos foi computado. Nesse contexto, para a construção de uma árvore filogenética baseada em máxima verossimilhança foi utilizado o modelo WAG (WHELAN; GOLDMAN, 2001) com distribuição gama com 2 categorias discretas. As análises partiram de uma árvore inicial que foi construída pelo método de *neighbor-joining*. Para avaliar a qualidade da árvore, foram executadas 1000 replicações de *bootstrap*. A árvore foi renderizada no software MEGA 11 (TAMURA; STECHER; KUMAR, 2021).

3.1.3.2 Filogenia entre as cistatinas de maqui e de *Oriza sativa*

Um alinhamento entre as cistatinas de maqui e de *Oriza sativa ssp japonica* (OC-I - P09229-2, OC-II - P20907, OC-III - Q6I570, OC-IV - Q5N806, OC-V - Q0JGM8, OC-VI - Q10Q46, OC-VII - Q10Q47, OC-VIII - Q10J94, OC-IX - Q10Q48, OC-X - P0C579, OC-XI - Q6K309, OC-XII - Q0JNR2; os códigos são referentes ao banco de dados UniProt) foi realizado na plataforma Mega 11 (TAMURA; STECHER; KUMAR, 2021) pelo método Muscle (EDGAR, 2004). A partir dele, uma árvore filogenética foi construída conforme descrito na seção 3.1.3.1.

3.2 AMPLIFICAÇÃO E CLONAGEM DAS ORFS

3.2.1 Isolamento das ORFs que codificam as cistatinas

As amostras de RNA total de *Aristotelia chilensis* foram submetidas ao processo de transcrição reversa para obtenção de cDNA com o kit High-Capacity cDNA Reverse Transcription (Applied Biosystems) e primers randômicos hexaméricos. O cDNA foi utilizado como molde na Reação em Cadeia da Polimerase (PCR) utilizando primers específicos (Tabela 1) contendo sítios para as enzimas de restrição *NdeI* e *Sall*.

Tabela 1: Sequências dos primers para amplificar as sequências codificantes das cistatinas de maqui. Os sítios de restrição *NdeI* (CATATG) e *Sall* (GTCGAC) estão destacados em negrito.

Cistatina	Primers
MaquiCPI-1	F: 5' ATTCATATGGGTGTGGAGGGTTTTGGGTC 3' R: 5' ATTGTCGACTTAATCGGTAGAAAGGTGCGAAATG 3'
MaquiCPI-2	F: 5' ATTCATATGTGGAAAGATCAAATGCCCGG 3' R: 5' ATTGTCGACTTATCCAATGACAACCTTAAAAGAAG 3'
MaquiCPI-3	F: 5' ATTCATATGAGGATAGATCCTCTTGTCGG 3' R: 5' ATTGTCGACTTACACAACCTTAAAAGAAGTAAGATT 3'
MaquiCPI-4	F: 5' ATTCATATGGCAAGAGTGGGTGGTGTAG 3' R: 5' ATTGTCGACCTAAGCAGTGGAGTCGGAAG 3'
MaquiCPI-5	F: 5' ATTCATATGGCCACCACCA CCACCGTCCG 3' R: 5' ATTGTCGACTCAGGAGTGGTCCGGCTCCA 3'
MaquiCPI-6	F: 5' ATTCATATGGGTTCGTTGCAACGAAGACTT 3' R: 5' ATTGTCGACCTAATTCATATAAACTTTCCTTC 3'

Para cada uma das sequências de cistatinas foi realizada uma PCR em um volume final de 25 μ L. Foram utilizados 200 μ M de cada dNTP (Promega), tampão Taq High Fidelity Pol Reaction Buffer 1X [Tris-HCl pH 8,8 (25°C), KCl, MgCl₂ (2,5 mM)] (Cellco, São Carlos), 0,5 μ M de cada *primer* e 2,5 U da enzima Taq High Fidelity Pol (Cellco, São Carlos). O protocolo de amplificação foi realizado no termociclador Veriti (Applied Biosystems) iniciando com uma temperatura de 95 °C por 3 minutos, seguido por 35 ciclos de amplificação de 30 segundos a 95 °C (desnaturação), 1 minuto na temperatura de hibridização (Tabela 2) e 1 minuto a 72 °C (elongamento) e terminando com uma extensão final de 7 minutos a 72 °C. Os *amplicons* foram analisados em gel de agarose 1% contendo brometo de etídio 1 μ g/mL.

Tabela 2: Temperaturas de hibridização utilizadas nas amplificações de cada uma das sequências codificantes das cistatinas de maqui.

MaquiCPI-1	MaquiCPI-2	MaquiCPI-3	MaquiCPI-4	MaquiCPI-5	MaquiCPI-6
47°C	55°C	47°C	50°C	45°C	40°C

3.2.2 Construção de vetores de expressão

Os produtos das amplificações descritos na seção 3.2.1 foram purificados com o Purification PCR kit (Cellco) e clonados no vetor de propagação pTZ57R/T (Fermentas). Para isso, foi utilizado o kit InsTAclone PCR Cloning Kit (Thermo Scientific), com as ligações ocorrendo em um volume final de 20 μ L, contendo 100 ng do plasmídeo, aproximadamente 35 ng (MaquiCPIs de 1 a 4) ou 65 ng (MaquiCPI-5) de inserto, 2,5 U da enzima T4 DNA ligase (Invitrogen) e tampão de ligação 1X [50 mM Tris-HCl (pH 7.6), 10 mM MgCl₂, 1 mM ATP, 1 mM DTT, 5% (m/v) polietilenoglicol-8000]. As reações de ligação foram conduzidas a 4 °C por 16 horas e então utilizadas para transformar, por choque térmico, células competentes de *Escherichia coli* DH5 α [F⁻ ϕ 80lacZ Δ M15 Δ (lacZYA-argF) U169 recA1 endA1 hsdR17(rK⁻, mK⁺) phoA supE44 λ ⁻ thi-1 gyrA96 relA1] preparadas por tratamento com cloreto de cálcio (SAMBROOK; RUSSEL, 2000). Os clones bacterianos contendo os vetores recombinantes (pTZ57R/T + inserto) foram selecionados em meio LB suplementado com carbenicilina (50 μ g/mL). Na superfície do meio foi espalhado 20 μ L de IPTG (Isopropil β -D-1-tiogalactopiranosídeo) (200mM) e 40 μ L de X-gal (5-bromo-4-cloro-3-indolil- β -D-galactopiranosídeo) (20 mg/mL), substrato da enzima β -galactosidase, para seleção azul e branco.

As colônias selecionadas foram avaliadas por PCR de colônia. As reações, de volume final 20 μ L, foram realizadas com tampão 1X Taq Buffer [KCl 10 mM Tris-HCl (pH 8.8 a 25

°C), 50 mM KCl, 0,08% (v/v) Nonidet P40] (Thermo Scientific), 4 mM de MgCl₂, 200 mM de cada dNTP, 0,3 μM de cada um dos primers M13 *Forward* (5' GTAAAACGACGGCCAG 3') e *Reverse* (5' GAGGAAACAGCTATGACCATG 3') e 1,25 U da enzima Taq Polimerase (Thermo Scientific). As colônias brancas foram retiradas com palitos de dente estéreis e utilizadas como molde da reação. O protocolo de amplificação foi realizado no termociclador Veriti (Applied Biosystems) iniciando com uma temperatura de 95 °C por 3 min, seguido por 35 ciclos de amplificação de 30 s a 95 °C, 1 min a 50 °C e 1 min a 72 °C e terminando com uma extensão de 7 min a 72 °C. Os *amplicons* foram analisados em gel de agarose 1%.

Uma das colônias positivas foi repicada em 5 mL de meio LB com 50 μg/mL de carbenicilina, e cultivadas por 16 horas a 250 rpm. O DNA plasmidial foi extraído com o kit QIAprep Spin Miniprep (QIAGEN). Os plasmídeos obtidos foram restringidos com as enzimas *NdeI* e *Sall*, os fragmentos referentes a cada maquiCPI foram isolados em gel e subclonados no vetor pET-28a (Novagen) previamente digerido com as mesmas enzimas. Nas digestões para liberação dos fragmentos, assim como do vetor pET-28a, foram utilizados cerca de 2 μg de DNA plasmidial em reações de 30 μL. Essas foram conduzidas em tampão Orange [Tris-HCl (50 mM) (pH 7,5 a 37°C), MgCl₂ (10 mM), NaCl (100 mM), BSA (0,1 mg/ml)], com 5 U das enzimas *NdeI* e *Sall*, por 16 horas a 37 °C. Todos os produtos de clivagem foram submetidos à eletroforese em gel de agarose 1%. As bandas de interesse foram excisadas do gel e purificadas com o kit QIAquick Gel Extraction Kit (QIAGEN).

As reações de ligação foram realizadas em volume final de 20 μL com 100 ng de vetor, aproximadamente 20 ng (MaquiCPIs de 1 a 4) e 35 ng (MaquiCPI-5) de inserto, 2,5 U da enzima T4 DNA ligase (Invitrogen) e tampão de ligação 1X. Essas foram conduzidas a 4 °C *overnight* e, por choque térmico, foram utilizadas para transformar células de *E. coli* DH5α competentes por CaCl₂. Os clones bacterianos contendo os vetores recombinantes (pET + inserto) foram selecionados em meio LB suplementado com canamicina (25 μg/mL) por 16 horas a 37 °C. Algumas colônias foram então selecionadas e avaliadas por PCR de colônia com os primers T7prom (5' TAATACGACTCACTATAGGG 3') e T7term (5' TAGTTATTGCTCAGCGG 3') e clivagem com as enzimas utilizadas na clonagem.

3.3 EXPRESSÃO DAS CISTATINAS RECOMBINANTES

3.3.1 Expressão das cistatinas recombinantes em células de *E. coli*

As maquicistatinas recombinantes foram expressas na cepa *E. coli* Rosetta (DE3) [F-ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)]. Os plasmídeos construídos foram utilizados para a transformação por choque térmico em células quimiocompetentes que, após transformação, foram plaqueadas em meio LB ágar com 25 µL/mL de canamicina e cloranfenicol, e colocadas em estufa a 37 °C overnight.

Uma das colônias foi repicada em um pré-inóculo de 5 mL de meio LB líquido contendo 25 µL/mL de canamicina e cloranfenicol, que foi mantido por 16 horas a 250 rpm e 37 °C. Após esse período, foi realizada uma diluição de 1:100 de pré-inóculo em 500 mL de LB líquido com 25 µL/mL dos mesmos antibióticos, e a cultura foi mantida nas mesmas condições de temperatura e agitação até chegar em sua fase logarítmica de crescimento (DO = 0,5 e $\lambda = 600$ nm). A produção de proteínas foi induzida pela adição de IPTG (isopropil β -D-1-tiogalactopiranosídeo) para concentração final de 0,4 mM. Alíquotas de 500 µL da cultura foram coletadas em intervalos de 1 hora para análise da expressão em SDS-PAGE 15% (LAEMMLI, 1970).

3.3.2 Lise bacteriana e teste de solubilidade

Após o período de indução, o volume total da cultura foi submetido à centrifugação em Sorvall RC5C plus rotor SLA15000 (4450 rpm; 4 °C; 10 min). O precipitado contendo as bactérias foi ressuspendido em 50 mL de tampão de lise (Tris 10 mM, NaH₂PO₄ 50 mM e NaCl 300 mM, pH 8) por litro de cultura e submetido à sonicação com 6 pulsos de 1 minuto com 20% de potência, intercalados por pausas de 30 segundos, em sonicador Sonic Desmembrator 500 (Fischer Scientific), para lise celular. O lisado foi centrifugado em Sorvall RC5C plus rotor SS34 (13000 rpm; 4 °C; 10 min), para separar as frações solúvel (sobrenadante) e insolúvel (precipitado). A fração solúvel foi filtrada a vácuo em filtro de 0,45 µm e a fração insolúvel foi ressuspendida em tampão de lise (50 mL por litro de cultura). Alíquotas de ambas foram recolhidas para análise em SDS-PAGE 15%.

Análises complementares a respeito da solubilidade das proteínas em *Escherichia coli* foram realizadas utilizando o *software* Soluprot (HON *et al.*, 2021).

3.3.3 Purificação das cistatinas recombinantes

As cistatinas foram purificadas a partir das frações solúveis dos sonicados, por cromatografia de afinidade em coluna contendo 5 mL de resina Ni-NTA-agarose (QIAGEN). Inicialmente, cada coluna foi pré-equilibrada com 5 volumes de tampão de lise (25 mL). Uma vez que a fração solúvel tenha passado a coluna, a resina foi lavada com 3 volumes do tampão (15 mL). Depois disso, as proteínas foram eluídas com 2 volumes (10 mL) desse tampão contendo concentrações crescentes de imidazol (10 mM, 25 mM, 50 mM, 75 mM, 100 mM) e com 3 volumes (15 mL) da fração mais concentrada (250 mM). As frações eluídas foram analisadas por SDS-PAGE 15%.

As frações purificadas foram dialisadas utilizando membrana SnakeSkin Dialysis Tubing (ThermoFischer Scientific) de 3500 MWCO contra 2 litros de tampão PBS (NaCl 1,37 mM; KCl 2,7 mM; Na₂HPO₄ 10 mM; NaH₂PO₄ 2 mM) pH 7,4 (MaquiCPIs 1, 2 e 5) ou pH 8 (MaquiCPIs 3 e 4) a 4 °C. Foram realizadas três trocas de tampão de 1 hora. As proteínas foram filtradas em filtro Millex® GV de 22 µm/ 25 mm (Millipore), quantificadas com o kit Pierce™ BCA Protein Assay e armazenadas a -20 °C em tubos tipo Falcon, até o uso.

3.4 INIBIÇÃO DA ATIVIDADE DE CISTEÍNO PEPTIDASES UTILIZANDO MAQUICISTATINAS

A caracterização bioquímica das cistatinas de maqui foi realizada a partir de ensaios de inibição da atividade de cisteíno peptidases. Para isso, foram utilizadas as enzimas papaína (10 nM) de *Carica papaya* (Calbiochem), além da catepsina L (Sigma-Aldrich) (3,5 nM) e da catepsina B (Calbiochem) (1,9 nM) de *Homo sapiens*.

Os experimentos foram conduzidos em cubetas de quartzo contendo tampão acetato de sódio 100 mM pH 5,5 a 37 °C em um volume final de 500 µL. Primeiramente, as peptidases foram encubadas no tampão contendo o ativador 2,5 mM de DTT (ditiotretitol) durante 5 minutos. Após esse período, o substrato fluorogênico Z-Phe-Arg-AMC (carbobenzoicarbonil-Phe-Arg-7-amino-4-metilcoumarina) (2,5 µM, 8 µM e 20 µM para papaína e catepsinas L e B, respectivamente) foi adicionado ao sistema e a atividade foi avaliada a partir da sua taxa de hidrólise associada ao monitoramento da fluorescência liberada a partir do AMC (excitação 380 nm e emissão 440 nm) em espectrofluorímetro Hitachi F-2500 (Figura 4A). A cada minuto, foram adicionadas concentrações crescentes de cistatina, sem alterações significativas no

volume da reação. A cada adição foram feitas leituras das atividades residuais. Os experimentos foram realizados em triplicata.

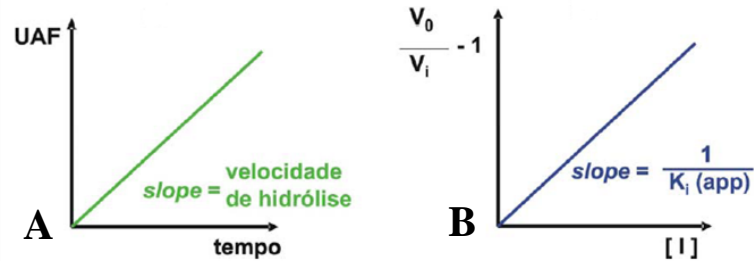


Figura 4: Determinação da atividade enzimática (A) e da constante de inibição (K_i) (B)

A partir disso, os valores obtidos das velocidades de hidrólise na ausência e na presença dos inibidores em diferentes concentrações foram utilizados para construir um gráfico (Figura 4B), cuja inclinação indica o inverso da constante de inibição aparente, como descreve a equação (1) (NAGASE; SALVESEN, 2001).

$$\frac{V_0}{V_i} - 1 = \frac{[I]}{K_{i_{app}}} \quad (1)$$

Onde:

V_0 = Velocidade de hidrólise na ausência de inibidor

V_i = Velocidade de hidrólise na presença de inibidor

$[I]$ = Concentração de inibidor

$K_{i_{app}}$ = Constante de inibição aparente

O valor de K_i pode ser então obtido de acordo com a equação (2) (NAGASE; SALVESEN, 2001). Para isso, os seguintes valores de K_m foram utilizados a fim de corrigir o valor de K_i para competição pelo substrato: 23 μM catepsina B, 2 μM catepsina L e 10 μM papaína (MELO *et al.*, 2001).

$$K_i = \frac{K_{i_{app}}}{1 + \frac{[S]}{K_m}} \quad (2)$$

Onde:

K_m = Concentração de substrato que gera o valor numérico da metade da velocidade máxima

$[S]$ = Concentração de substrato

K_i = Constante de inibição

4 RESULTADOS E DISCUSSÃO

4.1 IDENTIFICAÇÃO DAS SEQUÊNCIAS E ANÁLISE *IN SILICO*

O RNA total de plantas inteiras de maqui mantidas em TIB foi sequenciado com sucesso. A amostra cultivada sem ABA, coletada no início do experimento, gerou 76776826 *reads* e 72541 *contigs*. Dentre estes, a partir de uma investigação por BLASTP utilizando sequências de cistatinas de grupos e espécies vegetais distintas, foram identificadas seis prováveis sequências de diferentes cistatinas de *Aristotelia chilensis*. As sequências de nucleotídeos de cada uma das cistatinas de maqui são mostradas nas figuras de 5 a 10, com a sua respectiva tradução e indicação dos motivos conservados. A Tabela 3 mostra o tamanho das ORFs e as massas moleculares teóricas das cistatinas por elas codificadas.

MaquiCPI-1

```

-----UTR 5'-----
atgcaccaccttaatttgcggtgtccaaacaccacccccattctataaataaataataaccattctcctcattaatattcacaaaatc
----- M V K P L V I L S V C L I V
taaatctaaatctaaatcaaaagccatggtaaaa ATG GTA AAA CCT CTA GTA ATA TTG TCA GTT TGT TTG ATT GTA

S I C F C G V E G F G S I V G G R T E V K D
TCA ATT TGT TTC TGT GGT GTG GAG GGT TTT GGG TCA ATA GTG GGA GGG AGA ACG GAG GTG AAA GAT

V K T N E E V Q E L G R F S V E E F N R N Q
GTT AAG ACC AAT GAG GAG GTG CAA GAA CTG GGG AGG TTT TCG GTA GAG GAA TTC AAT AGA AAC CAG

I H R M S S N G G Q L L M F S Q V I E A Q K
ATT CAC CGC ATG AGC AGC AAT GGT GGC CAA TTG TTG ATG TTC TCT CAG GTT ATA GAG GCG CAA AAG

Q V V S G I K Y F L K I Q A M K N G Q T R L
CAG GTG GTT TCT GGG ATC AAG TAT TTT CTC AAG ATT CAA GCC ATG AAA AAT GGA CAA ACG CGT TTG

F D S V V V V K P W I H N S K E L L H F A P
TTT GAT TCG GTG GTT GTG GTT AAG CCT TGG ATT CAT AAT TCC AAG GAG TTG CTT CAT TTC GCA CCT

S T D #
TCT ACC GAT TAA

atattagattagattaaggttaattttatggtatgtcaatgtaaatctgtggttaggtatggtggtgttgtattaatctcctgttgtt
ttgcttctttgtaaatgtttcctacttttggtttttgtgttggaatgtggttgaaccatgtttatgtgtcctcgaacatgtaaggg
aataaataaacattaaattttaaaaaaaaaa
-----UTR 3'-----

```

Figura 5: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-1. A região codificante da proteína é mostrada a partir do códon de iniciação ATG e se encerra no códon de terminação TAA, representado por “#”. Os resíduos destacados em amarelo representam as regiões conservadas em fitocistatinas. Os nucleotídeos em letra minúscula, antes do códon de iniciação e depois do códon de terminação representam as regiões UTR 5’ e 3’, respectivamente.

MaquiCPI-2

```

-----UTR 5'----- M  A  T  S  V  Q  P  V  L  P  Q  L  L  D  Y  K  A  E
agatcaaacaccccat ATG GCT ACT AGC GTG CAA CCA GTT TTG CCC CAA TTG TTG GAT TAT AAG GCA GAG

Q  P  L  L  F  N  G  A  I  S  A  W  K  D  Q  M  P  G  G  W  T  L
CAA CCT CTT CTG TTT AAT GGA GCT ATA AGT GCA TGG AAA GAT CAA ATG CCC GGT GGG TGG ACC CTA

I  K  D  V  S  K  P  H  V  K  E  I  G  E  F  A  V  D  E  Y  N  K
ATA AAG GAC GTG AGC AAG CCA CAC GTG AAA GAG ATC GGT GAG TTT GCA GTT GAT GAG TAT AAT AAG

R  S  K  A  A  L  K  F  K  T  V  I  C  G  W  T  Q  V  V  S  G  I
AGA TCA AAG GCT GCG TTG AAG TTT AAG ACC GTA ATT TGT GGA TGG ACT CAG GTG GTT TCT GGG ATC

N  Y  K  L  I  V  E  A  K  D  G  A  D  T  N  Y  Y  E  A  V  V  W
AAT TAT AAG CTT ATT GTT GAA GCT AAA GAT GGA GCG GAT ACA AAT TAT TAT GAA GCT GTA GTG TGG

E  K  A  W  Q  K  F  K  E  L  T  S  F  K  V  V  I  G  #
GAG AAA GCT TGG CAG AAG TTT AAG GAA CTT ACT TCT TTT AAG GTT GTC ATT GGA TAA

tgacataaactgggccagcagtgagcatgccagatgtcaccttagaagcgtataactaaataaatggcttgggtggaatgcctccttag
ttaaagatgTTTTatgtatgcatgcatacgttggagatgaagataataagctcgttgggtgtcaaggtttcagcctgtggctcggcat
ttatgcatgtatacattgataaatgtgtacttgttaagtactgtattattatgttaatgaaacccttcttcggatggatagcactg
gaattaaatatttatgttgttgagtggaaggtgata
-----UTR 3'-----

```

Figura 6: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-2. A região codificante da proteína é mostrada a partir do códon de iniciação ATG e se encerra no códon de terminação TAA, representado por “#”. Os resíduos destacados em amarelo representam as regiões conservadas em fitocistatinas. Os nucleotídeos em letra minúscula, antes do códon de iniciação e depois do códon de terminação representam as regiões UTR 5’ e 3’, respectivamente.

MaquiCPI-3

```

-----UTR 5'-----M K S L
ctatataaatccactcaccaccaccacaacccaaaaccgcaacctaataaaaaaacactccaaaaaacaacagttaca ATG AAA TCA CTT

T L L L V I S S L F L S L F A A S A R I D P
ACC CTT CTT CTA GTC ATT TCG TCA CTA TTT CTC TCT CTC TTC GCT GCC AGT GCA AGG ATA GAT CCT

L V G G W T P I T N M E D P H V K E I G E F
CTT GTC GGT GGG TGG ACC CCA ATA ACG AAC ATG GAA GAC CCG CAC GTG AAA GAG ATT GGT GAG TTT

A V D A Y N K G S K A A L K F E T V I S G E
GCA GTT GAT GCG TAC AAC AAG GGT TCA AAG GCT GCG TTG AAG TTT GAG ACC GTA ATT AGC GGA GAG

T Q V V S G T N Y K L I V A A K D G E A V S
ACT CAA GTG GTT TCT GGC ACC AAC TAT AAG CTT ATT GTG GCA GCT AAA GAT GGA GAA GCG GTT TCA

K N Y E A V V W E K P W V N F K N L T S F K
AAG AAT TAT GAG GCT GTC GTG TGG GAG AAA CCA TGG GTG AAT TTC AAG AAT CTT ACT TCT TTT AAA

V V #
GTT GTG TAA

tatggtgatctttttagtagtaatctaagtattatataaaagaattgagaataataaga
-----UTR 3'-----

```

Figura 7: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-3. A região codificante da proteína é mostrada a partir do códon de iniciação ATG e se encerra no códon de terminação TAA, representado por “#”. Os resíduos destacados em amarelo representam as regiões conservadas em fitocistatinas. Os nucleotídeos em letra minúscula, antes do códon de iniciação e depois do códon de terminação representam as regiões UTR 5’ e 3’, respectivamente.

MaquiCPI-4

```

-----UTR 5'-----
gactaaaagactttaacacgtaaaaaattgctttttgtcttttactctctctctaataatcaaatatggttgatataaaaaacagaggc
-----
acaggctcgcaattggcattgaggaagagaaaaggaccaaagaataaaaagagaaaaaa  M  A  R  V  G  G  V
ATG GCA AGA GTG GGT GGT GTT

S  D  V  E  G  S  A  N  S  L  E  I  D  S  L  A  R  F  A  V  D  E
AGC GAT GTT GAG GGA AGT GCC AAC AGT CTG GAG ATT GAT AGC CTC GCT CGT TTC GCT GTT GAT GAG

H  N  K  K  Q  N  A  L  L  E  F  K  K  A  V  N  V  K  K  Q  V  V
CAC AAC AAG AAA CAG AAT GCG CTT CTG GAG TTT AAG AAA GCG GTG AAT GTG AAA AAG CAA GTG GTT

S  G  T  I  Y  Y  I  T  L  E  A  I  D  G  G  Q  K  K  V  Y  E  A
TCT GGT ACC ATA TAC TAC ATA ACC TTG GAG GCA ATC GAT GGT GGT CAG AAG AAA GTT TAT GAA GCC

K  V  W  D  K  P  W  M  N  F  K  E  L  Q  E  F  K  L  V  G  D  A
AAG GTC TGG GAT AAG CCC TGG ATG AAC TTC AAG GAG TTG CAG GAG TTC AAG CTT GTT GGT GAT GCC

P  S  D  S  T  A  #
CCT TCC GAC TCC ACT GCT TAG

tcccttggttccagttgaaaaagaagcgtatgagaagggtgagacttgtctgtttgaataaatgtaacaaaattatcttgaattt
ggattttaaccttacaatattgtttagaatgtgtggattcatcttttctttcatggtacaattctcaagcttctttagaacagtg
gtaccgtgaagatgtttttgatataatgtttgtctgttttcacatatttaactgtactacctgaaatgcacatgctttgtatgtgaag
atctgttcaacagaaaaacta
-----UTR 3'-----

```

Figura 8: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-4. A região codificante da proteína é mostrada a partir do códon de iniciação ATG e se encerra no códon de terminação TAG, representado por “#”. Os resíduos destacados em amarelo representam as regiões conservadas em fitocistatinas. Os nucleotídeos em letra minúscula, antes do códon de iniciação e depois do códon de terminação representam as regiões UTR 5’ e 3’, respectivamente.

MaquiCPI-5

```

-----UTR 5'-----
aataaataaataaagagtgtcgagatacgaattctctctcttaattatcctactacactcttctgcgcttttgacaatttcacttgtct
----- M A T T T T V G G L H E
tctctgcttttgatatttacgatttgcaacaagaa ATG GCC ACC ACC ACC ACC GTC GGA GGA TTG CAC GAA

S Q G T Q N S V E I D G L A R F A V E E H N
TCC CAA GGA ACT CAG AAC AGT GTC GAG ATC GAC GGC CTC GCT CGT TTC GCG GTT GAA GAA CAT AAC

K K E N R I L E F A R V V K A Q E Q V V A G
AAG AAA GAG AAT AGA ATC CTG GAG TTT GCA AGG GTA GTA AAG GCG CAA GAA CAG GTA GTT GCT GGTA

T L H H L T V E V I D A G K K K L Y E A K V
ACT TTG CAT CAT CTT ACG GTC GAG GTT ATT GAC GCG GGT AAG AAG AAG CTC TAT GAA GCA AAG GTT

W V K P W L N F K E L Q E F K H V G D V P S
TGG GTG AAG CCG TGG TTG AAC TTT AAA GAA TTG CAA GAA TTC AAG CAC GTT GGT GAT GTC CCA TCT

F T S S D L G V K K D G H G P G W Q S V P A
TTT ACT TCT TCA GAT CTT GGT GTT AAG AAA GAT GGC CAT GGC CCT GGA TGG CAA TCT GTG CCA GCA

H D P E V Q G A A N H A I K T I Q Q R S N S
Cat gat cct gaa gtc cag ggt gca gca aat cat gct atc aag acc atc cag cag aggt cca act cat

L Y P Y E L Q E I V H A N A E V I E D S A K
TTG TAT CCT TAC GAA CTT CAA GAA ATC GTC CAT GCA AAC GCT GAG GTG ATA GAA GAT TCT GCA AAA

F D M L L K V K R G D K E E K L K V E V H K
TTT GAT ATG CTT CTC AAG GTC AAG AGG GGA GAT AAA GAA GAG AAG CTC AAG GTT GAG GTG CAC AAG

N N E G T F H L N Q M E P D H S #
AAC AAT GAA GGT ACT TTC CAT CTG AAT CAG ATG GAG CCG GAC CAC TCC TGA

tctcacagattgtgtgatgtgtctagatttctatgtgccaagtatcctatagctcttttgtcgcaacgaacttttatctgtgctgtaa
tgcttcgcttatacaaatgtgtgtataagacttgtgaatcactatttacataatgctagtcaaacacacagtcacacacacagcgtt
tcaggtgaaatgcttgaatatggttatttcaaaaatcacaattgtccttgcttggtagtaataattattctgtgagatcgaaagagcgt
-----UTR 3'-----

```

Figura 9: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-5. A região codificante da proteína é mostrada a partir do códon de iniciação ATG e se encerra no códon de terminação TGA, representado por “#”. Os resíduos destacados em amarelo representam as regiões conservadas em fitocistatinas. Nota-se a sequência SNSL, característica da inibição de legumainas. Os nucleotídeos em letra minúscula, antes do códon de iniciação e depois do códon de terminação representam as regiões UTR 5’ e 3’, respectivamente.

MaquiCPI-6

```

-----UTR 5'-----
tctccaagcaaaatcagaggaagatatctctatttcattctaatctttttgttctctcaagtttggtgacacgaggcggaagctggt
-----
ttcgccacgtcaaaaacgaattgattgatctgttcaaccctttgccttaccactaagtaaccaactaacattcatcgaccaaccgaca
-----
cgcttcaagtaaccaaagaatccaaataatagtcagtcctgttgggttgtaaaaagaacaaaagtgtaaaaagtgtaaaaagtgta
-----
-----M N K L S V I I L S V V V L V C G F
tttacagaggaag ATG AAT AAG TTG TCT GTG ATA ATC TTA AGT GTT GTA GTT TTG GTT TGT GGG TTT
V E L G R C N E D L I K M K L G G V H D C K
GTT GAA TTG GGT CGT TGC AAC GAA GAC TTG ATA AAG ATG AAA CTT GGA GGC GTT CAC GAT TGC AAA
G T Q N S A E I E S L A R F A V Q E H N N K
GGC ACT CAA AAC AGC GCC GAG ATC GAA AGC CTC GCT CGT TTC GCT GTC CAA GAA CAC AAC AAC AAA
E N S L L E F A R V L K A K E Q V V A G T M
GAG AAT TCA CTT CTT GAG TTT GCA AGG GTG TTG AAG GCC AAA GAA CAA GTT GTT GCT GGT ACG ATG
Y H L T L E A I D A G K K K I Y E A K V W V
TAC CAT CTT ACA CTT GAA GCA ATT GAT GCT GGC AAG AAG AAG ATA TAT GAA GCT AAA GTC TGG GTT
K P W M N F K Q L Q E F K H A Q D G P S V T
AAG CCA TGG ATG AAC TTC AAG CAG TTG CAG GAA TTC AAA CAT GCT CAG GAT GGC CCT TCC GTC ACC
T S N L G V K P D G H G L G W Q A V S T N D
ACT TCA AAC CTC GGC GTT AAA CCA GAT GGC CAT GGA CTA GGA TGG CAA GCA GTG TCA ACA AAT GAT
P E V E D A A N H A V K S I Q Q R S N S L S
CCT GAG GTA GAA GAT GCA GCA AAC CAT GCT GTT AAA TCC ATC CAG CAG AGG TCC AAC TCA TTG TCA
P Y E L L E I L L A K A K V I E D Y A K F E
CCC TAT GAA CTT CTA GAG ATC CTT CTA GCT AAG GCC AAG GTC ATT GAA GAT TAT GCC AAA TTT GAA
L L L K V R R G I K E E Q F R V E V T K N N
TTG CTT CTG AAG GTG AGG AGG GGA ATT AAA GAG GAG CAG TTC AGA GTT GAA GTA ACT AAG AAT AAT
E G K F Y M N #
GAA GGA AAG TTT TAT ATG AAT TAG
aatgaaactactgcccactggattcagtaattaagtactatattaaaagcatctatggattgaggctgtgatctactacaatgtgatg
taaaataatgaataaacccactttaacgatatgcattatgtgtaaatatgcaaattgcaatgtttgctggctgacttttgctcgatct
aaatacagctacatcctccttatatattagatgtgtatattatataaacaggctatgttactctcaattcatatttttaattcatc
atggtgtcagatgatgctaca
-----UTR 3'-----

```

Figura 10: Sequência dos nucleotídeos traduzidos do cDNA que codifica a MaquiCPI-6. A região codificante da proteína é mostrada a partir do códon de iniciação ATG e se encerra no códon de terminação TAA, representado por “#”. Os resíduos destacados em amarelo representam as regiões conservadas em fitocistatinas. Nota-se a sequência SNSL, característica da inibição de leguminais. Os nucleotídeos em letra minúscula, antes do códon de iniciação e depois do códon de terminação representam as regiões UTR 5’ e 3’, respectivamente.

Tabela 3: Tamanho das ORFs e massas moleculares teóricas das cistatinas de maqui.

Cistatina	Nucleotídeos (pb)	Aminoácidos	Massa molecular Teórica (kDa)
MaquiCPI-1	384	128	12,21
MaquiCPI-2	375	125	13,98
MaquiCPI-3	351	117	12,68
MaquiCPI-4	303	101	11,06
MaquiCPI-5	612	204	22,79
MaquiCPI-6	672	224	25,07

A inferência do peptídeo sinal é descrita na Figura 11 a partir do programa SignalP 5.0 (ALMAGRO ARMENTEROS *et al.*, 2019). Os peptídeos-sinal, ou sequências-sinal, são sequências sinalizadoras de endereçamento celular. Elas são caracterizadas como uma extensão, geralmente N-terminal da proteína, de cerca de 15 a 60 aminoácidos, que são clivados por peptidases-sinal no retículo endoplasmático rugoso. Esse tipo de sequência é geralmente caracterizado por possuir três domínios, que a partir do N-terminal são: (i) uma região positivamente carregada; (ii) uma hidrofóbica e (iii) outra polar não carregada. O sítio de clivagem da peptidase sinal é localizado nessa última região (VON HEIJNE, 1985). Nesse sentido, a análise das sequências de aminoácidos nos permitiu observar que as cistatinas MaquiCPI-1, MaquiCPI-3 e MaquiCPI-6, possuem sequências características de peptídeos sinais em suas N-terminais, que indicam que essas proteínas são direcionadas à rota de secreção.

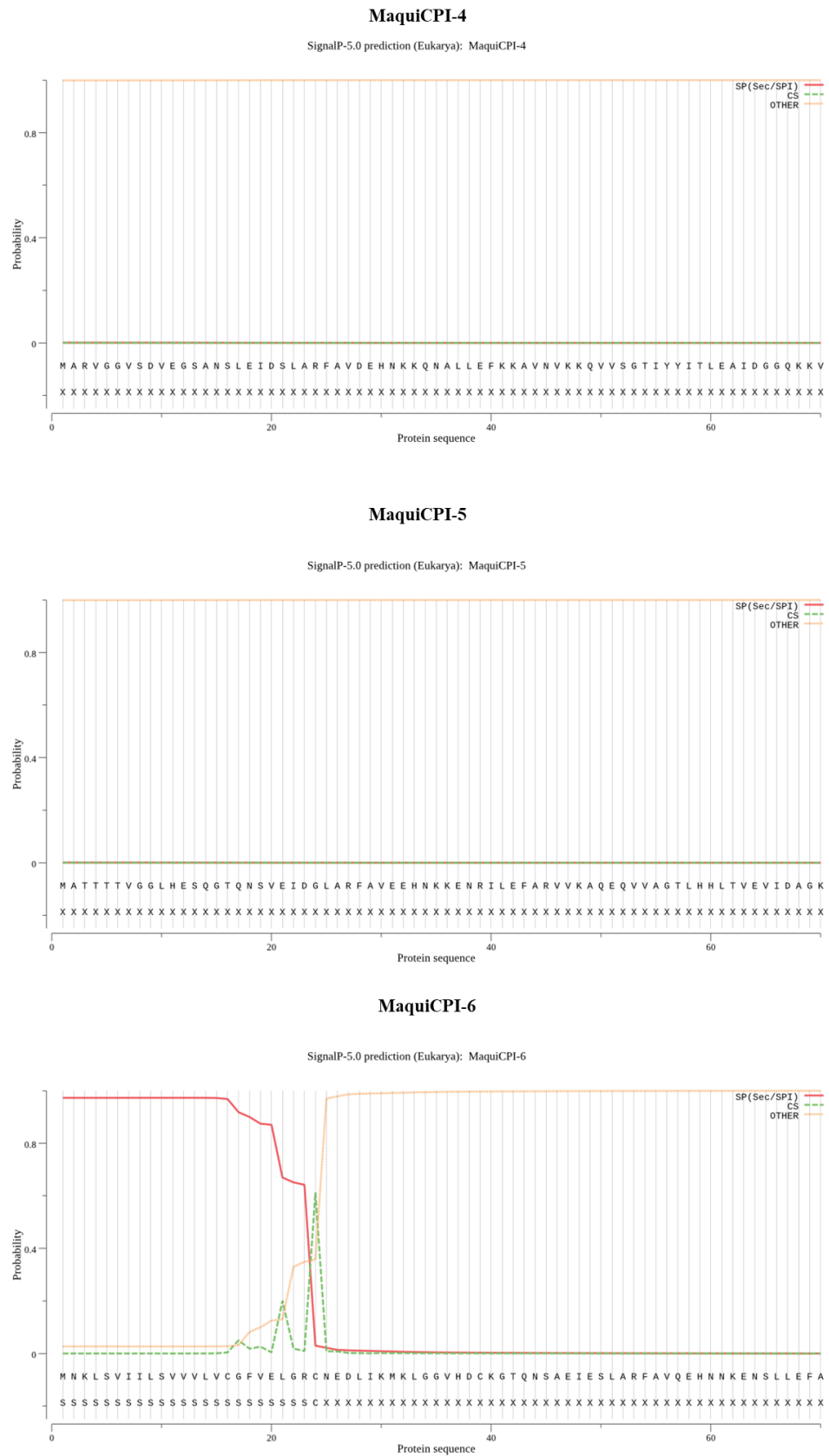
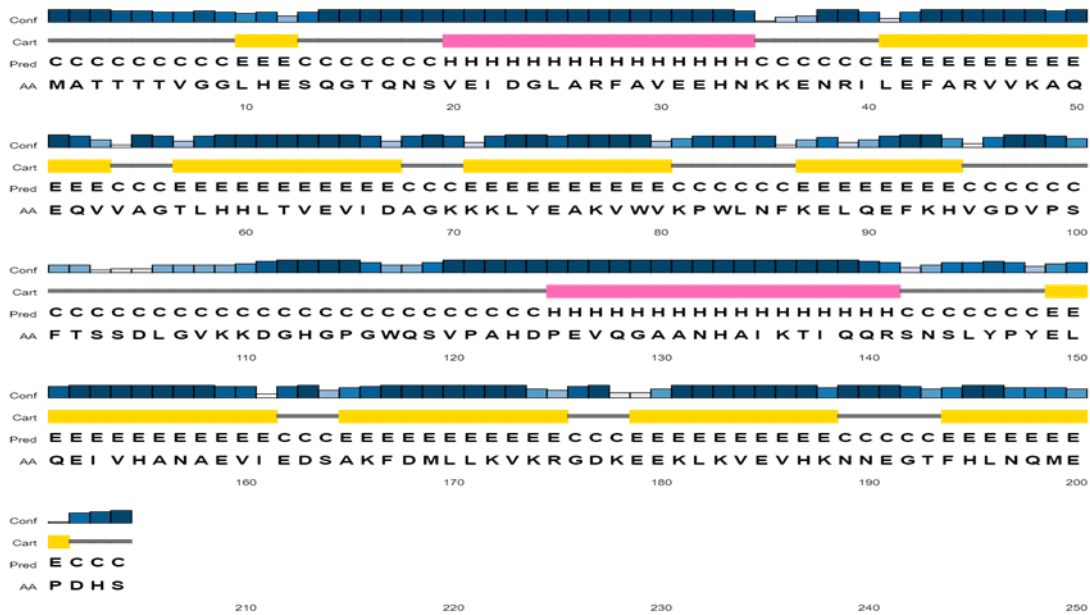


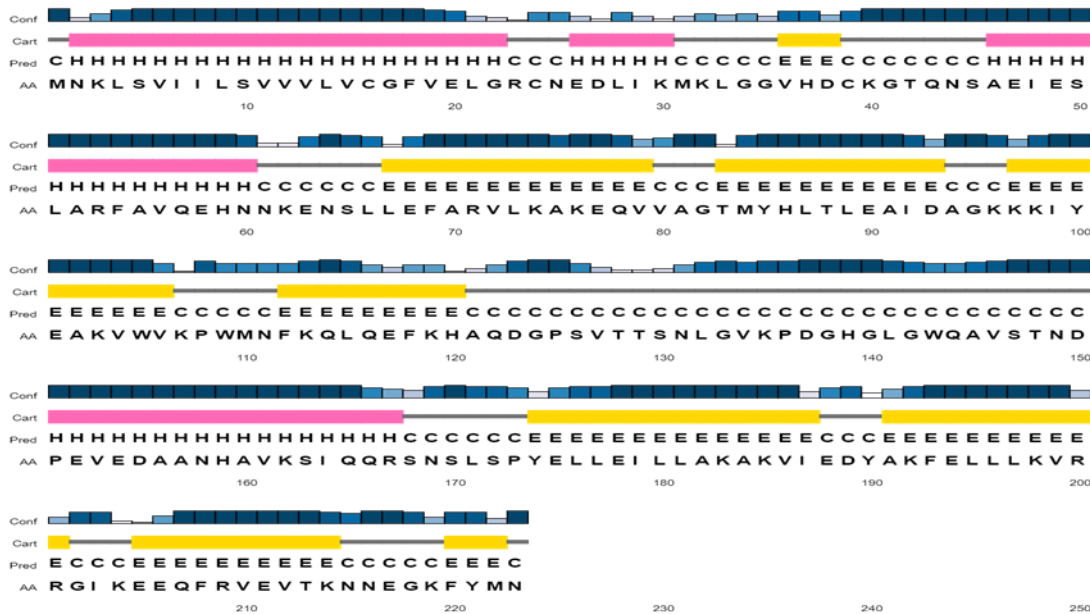
Figura 11: Predição de prováveis peptídeos sinais nas maquicistatinas. Os gráficos gerados pelo SignalIP 5.0 mostram a presença de seqüências com características de peptídeos sinal. Os prováveis sítios de clivagem estão indicados nos picos em verde.

A predição da estrutura secundária pelo programa PSIPRED (Figura 12) permitiu observar, como esperado, a presença de um resíduo de glicina na região N-terminal e de um *loop* contendo um triptofano na região C-terminal. O motivo QxVxG foi sempre observado entre uma folha β e um *loop* e a sequência LARFAVDEHN formou uma α -hélice. Ao analisar a região de peptídeo sinal, foi observada uma α -hélice característica deste tipo de sequência, não só nas MaquiCPIs 1,3 e 6, mas também na MaquiCPI-2. No entanto, a predição utilizando os programas SignalP, PrediSi e Phobious não indicaram um peptídeo sinal para a MaquiCPI-2. Por isso, sua sequência completa, incluindo esta região, foi clonada em pET28a e foram realizadas tentativas de produzir essa proteína em *E.coli*. No entanto, não houve sucesso na expressão. Esse fato levou à uma investigação mais profunda a respeito desta sequência. Uma análise utilizando o software de predição iPSORT, revelou, com alta probabilidade, que a sequência N-terminal, do aminoácido 1 ao 30 (um W), seria uma possível sinalizadora para o cloroplasto. No entanto, até o presente, não há relato na literatura de cistatinas nessa organela. Assim, seria interessante a realização de experimentos de localização subcelular em células vegetais, utilizando, por exemplo, a proteína inteira (ou apenas a sequência N-terminal) fusionada a uma proteína fluorescente. A proteína recombinante expressa poderá também ser utilizada para a produção de anticorpos e realização posterior de imunolocalização. É sabido que os peptídeos-sinais costumam dificultar a expressão de proteínas em *E. coli* (GIANOTTI *et al.*, 2006). Desta forma, ao retirar esta sequência, a expressão da proteína foi bem sucedida, o que também corrobora a suspeita de que ela deve estar envolvida com a localização da proteína na planta. Para estabelecer o tamanho da região N-terminal da MaquiCPI-2 a ser excluído, utilizamos o alinhamento com a MaquiCPI-3, mostrado na Figura 13. A predição de clivagem do peptídeo sinal para a MaquiCPI-3 é entre os resíduos 22 e 23 (AASA-RI). O alinhamento mostra uma região similar em MaquiCPI-2 (AISA-WK), entre os resíduos 29 e 30. Após seis resíduos desta região há um bloco conservado entre as duas proteínas (GGWT). Por isso, com base nestas similaridades e para manter as distâncias prováveis, excluimos a sequência entre os resíduos 29 e 30, mantendo o triptofano (W), que na predição do iPSORT, fazia parte do provável peptídeo sinalizador.

MaquiCPI-5



MaquiCPI-6



Legend:

- Fita β
- α-hélice
- coil

- Conf:** - Confiança da predição
- Cart:** - Predição da estrutura (Cartoon)
- Pred:** - Predição da estrutura
- AA:** - Sequência alvo

Figura 12: Predição da estrutura secundária das maquicistatinas. As fitas beta estão representadas pela cor amarela enquanto as alfa hélices pela cor rosa. Os quadrados em diferentes tonalidades de azul indicam a confiança da predição.

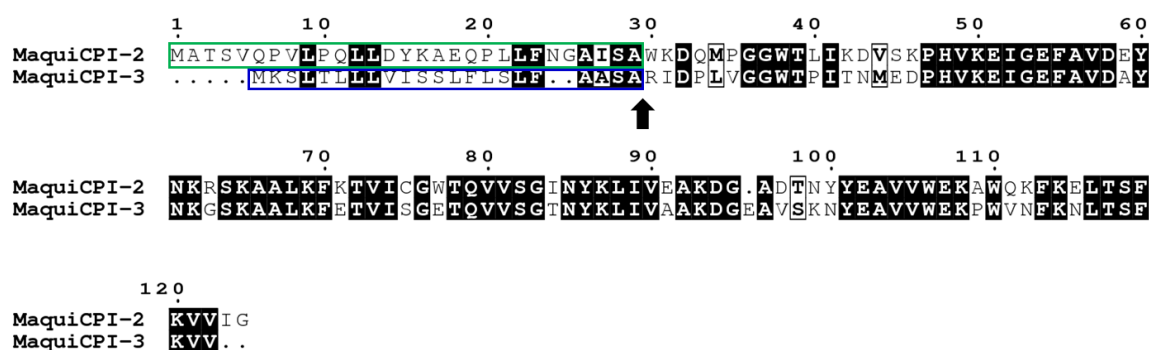


Figura 13: Alinhamento entre MaquiCPI-2 e MaquiCPI-3. O alinhamento foi realizado com o software Muscle. A caixa em azul indica a provável região de peptídeo sinal identificada pelo programa SignalIP, enquanto a caixa verde representa a possível sequência sinalizadora para o cloroplasto. A seta indica o provável sítio de clivagem dos peptídeos sinais nas duas proteínas, a partir de onde elas foram clonadas para expressão.

A Tabela 4 compara algumas características dessas proteínas com e sem sequência sinalizadora.

Tabela 4: Características das MaquiCPIs com ^(a) e sem ^(b) sequência sinalizadora.

		Número de aminoácidos	Massa Molecular Teórica (kDa)	pI
MaquiCPI-1	(a)	128	14,21	8,58
	(b)	109	12,15	8,27
MaquiCPI-2	(a)	125	13,98	8,59
	(b)	96	10,89	8,97
MaquiCPI-3	(a)	117	12,68	6,57
	(b)	95	10,37	5,80
MaquiCPI-4	(a)	-	-	-
	(b)	101	11,06	5,83
MaquiCPI-5	(a)	-	-	-
	(b)	204	22,76	6,06
MaquiCPI-6	(a)	224	25,07	7,71
	(b)	203	22,82	7,90

Para determinar a semelhança entre as sequências, foi realizada uma análise de similaridade utilizando o BLASTP (Tabela 5) e um alinhamento entre as cistatinas da espécie estudada. A Figura 14 apresenta esse alinhamento gerado pelo software Muscle (EDGAR, 2004) e editado com o programa ESPrpt (ROBERT; GOUET, 2014) destacando as prováveis

sequências sinalizadoras e as regiões conservadas de fitocistatinas. Os dados obtidos demonstram que embora nem sempre haja alta identidade entre a maior porção das cistatinas, ela ocorre nos motivos característicos da família. Estes, provavelmente, se mantêm conservados, pois estão estritamente relacionados com a função da proteína. Nas cistatinas, particularmente, a glicina da região N-terminal, o domínio QxVxG e triptofano C-terminal são importantes para a interação com a peptidase alvo (BARRETT, 1987). Nota-se que no domínio QxVxG há ocorrência de duas variantes: QVVSG (MaquiCPIs de 1 a 4) e QVVAG (MaquiCPIs 5 e 6). Essa variação é relevante, pois pode influenciar na atividade inibitória contra determinadas peptidases, como é o caso das catepsinas-B serem melhor inibidas por cistatinas com o motivo QVVSG (AUERSWALD et al., 1992). Foi observado também alto grau de conservação no motivo SNSL da MaquiCPI-5 e MaquiCPI-6. Essa sequência é de grande relevância para a interação com cisteíno peptidases do tipo legumaína, e respectiva inibição (MARTINEZ et al., 2007). A região conservada da alfa hélice, exclusiva das cistatinas de plantas, apresentou o seguinte consenso: [LI]-[AG]-[RE]-F-[AS]-V-[EDQ]-[EA]-[HYF]-N. Por fim, observa-se também, o que é mais evidente na Figura 13, o alto grau de similaridade entre a MaquiCPI-2 e MaquiCPI-3, o que possibilita inferir que uma possa ser derivada da outra dentro do processo evolutivo, sendo então direcionadas a compartimentos distintos. A MaquiCPI-1, por sua vez, difere de todas as outras por possuir aminoácidos a mais (8 resíduos) na região entre a alfa-hélice conservada e o motivo QxVxG.

Tabela 5: Similaridade entre as MaquiCPIs obtidas por BLASTP.

	CPI-1	CPI-2	CPI-3	CPI-4	CPI-5	CPI-6
CPI-1	100%	51%	53%	57%	61%	56%
CPI-2	51%	100%	76%	69%	54%	55%
CPI-3	53%	76%	100%	60%	56%	57%
CPI-4	57%	60%	60%	100%	84%	82%
CPI-5	61%	54%	56%	84%	100%	87%
CPI-6	56%	55%	57%	82%	87%	100%

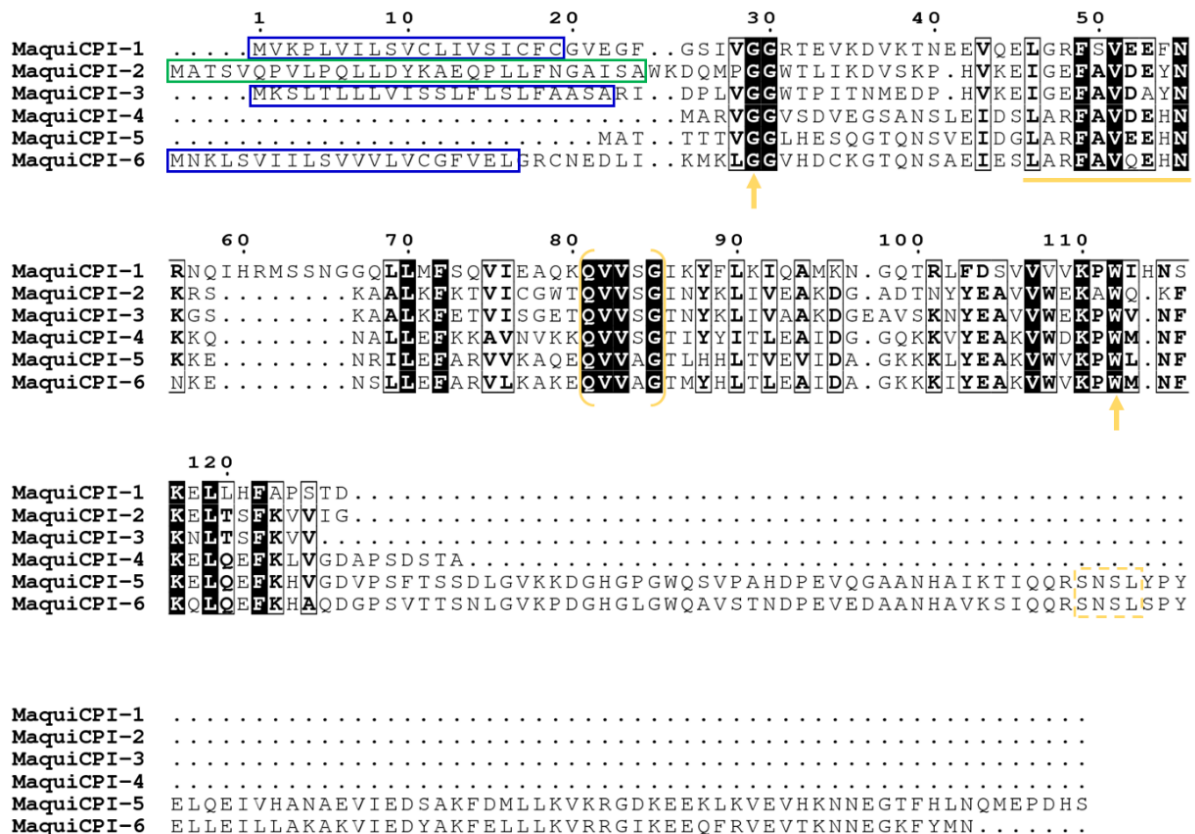


Figura 14: Alinhamento das seqüências de aminoácidos das MaquiCPIs pelo software Muscle. As caixas em azul indicam as regiões prováveis de peptídeo sinal identificadas pelo programa SignalIP, enquanto a caixa verde representa a seqüência sinalizadora para o cloroplasto. Destacados em amarelo, estão as regiões conservadas das fitocistatinas. A primeira seta indica a glicina N-terminal e a linha a seqüência [LVI]-[AGT]-[RKE]-[FY]-[AS]-[VI]-x-[EDQV]-[HYFQ]-N. Entre colchetes, está o domínio QxVxG. A segunda seta apresenta o triptofano C-terminal e a caixa pontilhada o domínio de inibição de legumaínas SNSL. Se o *score* de similaridade para uma coluna é maior que o valor de *threshold* (0,7 no programa ESPript), os resíduos considerados altamente similares são destacados em negrito e indicados com uma caixa vazia de contorno preto. Em caso de identidade estrita, a coluna é colorida de preto.

No banco de dados NCBI estão depositadas seqüências de cistatinas de várias espécies vegetais. Como esperado, algumas delas apresentam identidade com as cistatinas de maqui, conforme observado na Tabela 6 e nos alinhamentos da Figura 15 à Figura 20. As cistatinas do banco de dados que apresentaram a mais alta identidade com as MaquiCPIs 1, 2 e 3 compartilham 55% a 70% dos seus aminoácidos, o que é inferior ao observado para as MaquiCPIs 4, 5 e 6, cuja identidade está na faixa de 75% a 85%. Assim, nota-se que as MaquiCPIs 1, 2 e 3 possuem estrutura primária mais variável, o que pode refletir em diferenças relacionadas a aspectos funcionais. Em relação à similaridade, essa diferença é observada, porém menos expressiva. Os resultados do BLASTP apresentaram exemplares do clado das rosídeas (subdivisão das angiospermas), o qual o maqui pertence, mas houve poucas seqüências associadas a clados mais internos, provavelmente devido à escassez dessas no banco. Da família do maqui, a Elaeocarpaceae, por exemplo, não há qualquer exemplar de cistatina no NCBI. A

espécie mais próxima de *Aristotelia chilensis* recorrentemente apresentada na tabela, *Cephalotus follicularis*, está presente na mesma ordem, a Oxidales. Desse modo, as maquicistatinas são as primeiras a serem descritas entre as espécies da família Elaeocarpaceae.

Tabela 6: Identidade e similaridade da estrutura primária entre as maquicistatinas e outras espécies vegetais. Os códigos de acesso no banco de dados NCBI estão dispostos ao lado no nome da espécie.

Cistatina	Espécie Vegetal	Identidade (%)	Similaridade (%)
MaquiCPI-1	<i>Cephalotus follicularis</i> (GAV64583.1)	65,15	77
	<i>Theobroma cacao</i> (XP_007034159.1)	63,64	85
	<i>Durio zibethinus</i> (XP_022735006.1)	64,17	83
	<i>Trema orientale</i> (PON95535.1)	62,79	76
MaquiCPI-2	<i>Gossypium aridum</i> (MBA0694727.1)	65,62	77
	<i>Durio zibethinus</i> (XP_022721053.1)	60,42	76
	<i>Hibiscus syriacus</i> (XP_038995199.1)	55,24	73
	<i>Cephalotus follicularis</i> (GAV65231.1)	63,92	74
MaquiCPI-3	<i>Juglans regia</i> (XP_018838679.1)	68,37	84
	<i>Momordica charantia</i> (XP_022142402.1)	69,15	81
	<i>Cephalotus follicularis</i> (GAV65232.1)	69,70	78
	<i>Hibiscus syriacus</i> (XP_039060007.1)	59,65	77
MaquiCPI-4	<i>Hevea brasiliensis</i> (ACZ02398.1)	81,19	96
	<i>Humulus lupulus</i> (6VLP_A)	78,22	93
	<i>Salix dunnii</i> (KAF9675272.1)	77,23	91
	<i>Populus alba</i> (XP_034921130.1)	76,24	91
MaquiCPI-5	<i>Vitis vinifera</i> (XP_002267841.1)	82,50	90
	<i>Nyssa sinensis</i> (KAA8537359.1)	83,50	90
	<i>Quercus lobata</i> (XP_030966718.1)	79,80	89
	<i>Cephalotus follicularis</i> (GAV82623.1)	79,90	89
MaquiCPI-6	<i>Tripterygium wilfordii</i> (XP_038719056.1)	80,80	91
	<i>Pistacia vera</i> (XP_031272072.1)	79,02	89
	<i>Cephalotus follicularis</i> (GAV86994.1)	82,74	88
	<i>Morella rubra</i> (KAB1222999.1)	79,05	87

```

1          10          20          30          40          50          60
MaquiCPI-1      M V K P . . . L V I L S V C L I V S I C F C G . V E C F C S I V G C R T E V K D V K T N E E V O E L C R F S V E E F N R N Q I H R M S . . .
Cephalotus_follicularis M A K P S M L L I L L S C S V I A T L S M S G T V C F C S I V G C R T E V K D V E K N V E V O E L C R F S V E E F N R S R . . . . .
Theobroma_cacao      M A K A E V . . I L L S L L S V I T L S F T A . V N C Y C T L V G C R T E I I D V K T N K E V O E L C R F S V E E F N R S R . G G L R . . .
Durio_zibethinus     M A T T D V . . I K L L S L L V I A L S F T V . V N C Y C T M V G G R T E I I D V K T N E V O E L C R F A V E E F N R S Q . G R F S . . .
Trema_orientale      M A L I K . . . S P A V T L A I V L S L L F A S . A T C G E L V G G R T A I I D V K T N K E V O E L C R F S V E E Y N R S Q I R R G R R A G

          70          80          90          100         110         120
MaquiCPI-1      . . S N G G Q L L M F S Q V T E A O K Q V V S C I K Y Y L K I E A M . R N C Q T R L F D S V V V V K P W I H N S K E L L E F A P S T D . . . . .
Cephalotus_follicularis . . N G G A L L R F S R V V E A O K Q V V S C I K Y Y L K I E A T . R N G E I S F F D S V V V V K P W V H N S K E L L E F T P C V E Y . . . . .
Theobroma_cacao      . . S N G G T E L L V F S Q V V E A O K Q V V S C I K Y Y L K I E A M . Q S G L T R T F S V V V V K P W . E D S K E L L E F T P S P L . . . . .
Durio_zibethinus     . . S D G G V E L V F S Q V V E A O K R V V S C I K Y Y L K I E A T . L N G Q T K I T F D S V V V V K P S . D E S K E M L N E T P S T . . . . .
Trema_orientale      N D A C A G D L R F S E V V E A E R Q V V S C I K Y Y L K I E A V A R K T G E Q R I T F D S V V V V K P W L H D S K E L L E F A P S T D R D V H F

```

Figura 15: Alinhamento entre a MaquiCPI-1 e cistatinas de outras espécies vegetais. Foram alinhadas as cistatinas de *Cephalotus follicularis* (GAV64583.1), *Theobroma cacao* (XP_007034159.1), *Durio zibethinus* (XP_022735006.1), *Trema orientale* (PON95535.1) com a MaquiCPI-1 utilizando o programa Muscle.

```

1          10          20          30          40          50          60          70
MaquiCPI-2      M A T S V Q P V L F Q L L D Y K A E Q P L L F N G A T S A W K D O M P G G W T L K D V G R P H V E I G E F A V D E Y N K R S K A A L K F K
Gossypium_aridum    . . M Q O N S C F L I L S L I L L P L F F S P A R . . . K D A L A G G W T P K D I K D P H V E I A R F A V D E Y N K Q S K A S L K L L D
Durio_zibethinus    . . M H O N S R F L I L L S L L L P L I F S P A R . . . K D A L V G G W T P K N I K D P H V E I A R F A V D E Y N K Q S N T S L E L L
Hibiscus_syracus    . . M R R N L R F L I F L S L V F L P L L F S P A R . . . K A N L P G G W S P E D I T D P H V E I A R F A V D E Y N K H S K T S L E L E
Cephalotus_follicularis . . . M N P T L L V L V S L L L P L F A T S A R . . . T E A L A G G W T P K N V K D P H V E I G E F A V D E N N K N Q K V D L K F E

          80          90          100         110         120
MaquiCPI-2      T V I C K T O V V S G I N Y K L I V E A R D G E A T N Y E A V V W E K A W Q R E K E L T S F K V V I G
Gossypium_aridum    E V V R C E T O V V S G I N Y K L V L R A R D G E A V N Y E A V V W E K A W L H E R N L T S F T V V R G
Durio_zibethinus    R V V R C E T O V V A G I N Y R L V L N A T D G R A T N A Y Q A I V W E K A W Q N E K N L T S F N L V N G
Hibiscus_syracus    R V V R C E S O V V S G M N Y R L V L R A R D G T A V K T Y Q A V V W E K A W Q N E K S L T S F N L V K G
Cephalotus_follicularis R V I S C E S Q V V S G I N Y R L L T A R D G A V E N N Y E A V V W E K A W E R E K K L T S F K V V N C

```

Figura 16: Alinhamento entre a MaquiCPI-2 e cistatinas de outras espécies vegetais. Foram alinhadas as cistatinas de *Gossypium aridum* (MBA0694727.1), *Durio zibethinus* (XP_022721053.1), *Hibiscus syriacus* (XP_038995199.1) e *Cephalotus follicularis* (GAV65231.1) com a MaquiCPI-2, utilizando o programa Muscle

```

1          10          20          30          40          50          60
MaquiCPI-3      . . . . . M K S L T L L L V I S S L F L S L F . A A S A R I D P L V G G W T P I T N M E D P H V R E I G E F A V D A Y N K G S K A A L
Juglans_regia      . . . . . M K A Q F F Y L L T F L F L P I F S A A A G P R G P L L G G W O P I K N V K D P H V R E I G E F A V D E Y N K T S K A S L
Momordica_charantia M K S T L P M K S R S A P L L L L L L P A M V A A V . A V A A R K G P L V G G W T P I T D V K D P H V O E I G R F A V A A Y N T O S K G A I
Cephalotus_follicularis . . . . . M N P T L L V L V S L L L I L H L F . A T G G R T E A L V G G W T P I K N V K D P H V E I G Q F A V D E N N K N Q K A D L
Hibiscus_syracus    . . . . . M Q O N S R S L I F L E L S I V F E P L I . E S D A R I R A N E T G G W S P I Q N I T D P H V S E I A E F A V G Q Y N N E S K A S L

          70          80          90          100         110
MaquiCPI-3      R F E F V I S G E T Q V V S G T N Y K L I V A A R D G E A V S K N Y E A V V W E K P W N F K N L T S F K V V . .
Juglans_regia      R F Q N V V K G E T Q V V S G T N Y R L L V V A E D G . G V T K N Y Q A I V W E K P W E N F K L T S F T V P V K A
Momordica_charantia R F D R V E S G E T Q V V S G T N Y R L L V V A R D G G G S T A K Y Q A I V W E K P W E N F K L T S F R P V A N
Cephalotus_follicularis R F E R V I S G E S Q V V S G T N Y K L I L T A R D G . T V E N K Y E A V V W E K P W E R E K K L T S F K I V K A
Hibiscus_syracus    R L E R V L K G E T Q V V S G T N Y R L L V L K A R D G . A A E K T Y E A V V W E K P W Q N Y K N L T S F D I V K G

```

Figura 17: Alinhamento entre a MaquiCPI-3 e cistatinas de outras espécies vegetais. Foram alinhadas as cistatinas de *Juglans regia* (XP_018838679.1), *Momordica charantia* (XP_022142402.1), *Cephalotus follicularis* (GAV65232.1), *Hibiscus syriacus* (XP_039060007.1) com a MaquiCPI-3, utilizando o programa Muscle.

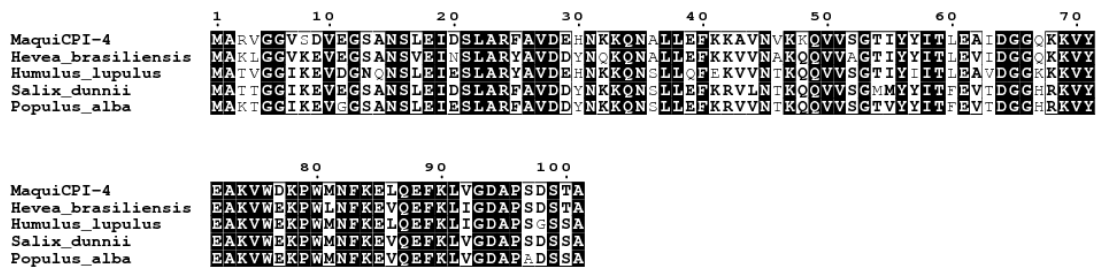


Figura 18: Alinhamento entre a MaquiCPI-4 e cistatinas de outras espécies vegetais. Foram alinhadas cistatinas de *Hevea brasiliensis* (ACZ02398.1), *Humulus lupulus* (6VLP_A), *Salix dunnii* (KAF9675272.1), *Populus alba* (XP_034921130.1) com a MaquiCPI-4, utilizando o programa Muscle.

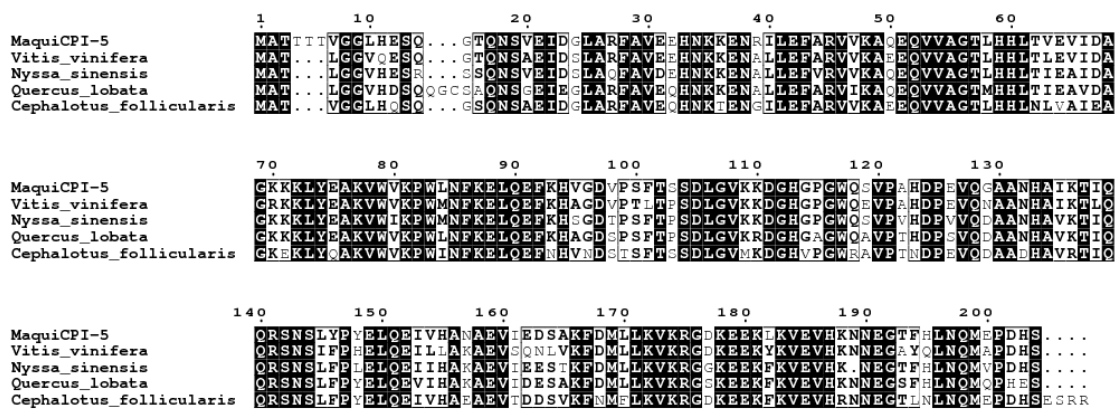


Figura 19: Alinhamento entre a MaquiCPI-5 e cistatinas de outras espécies vegetais. Foram alinhadas cistatinas de *Vitis vinifera* (XP_002267841.1), *Nyssa sinensis* (KAA8537359.1), *Quercus lobata* (XP_030966718.1), *Cephalotus follicularis* (GAV82623.1) com a MaquiCPI-5, utilizando o programa Muscle.

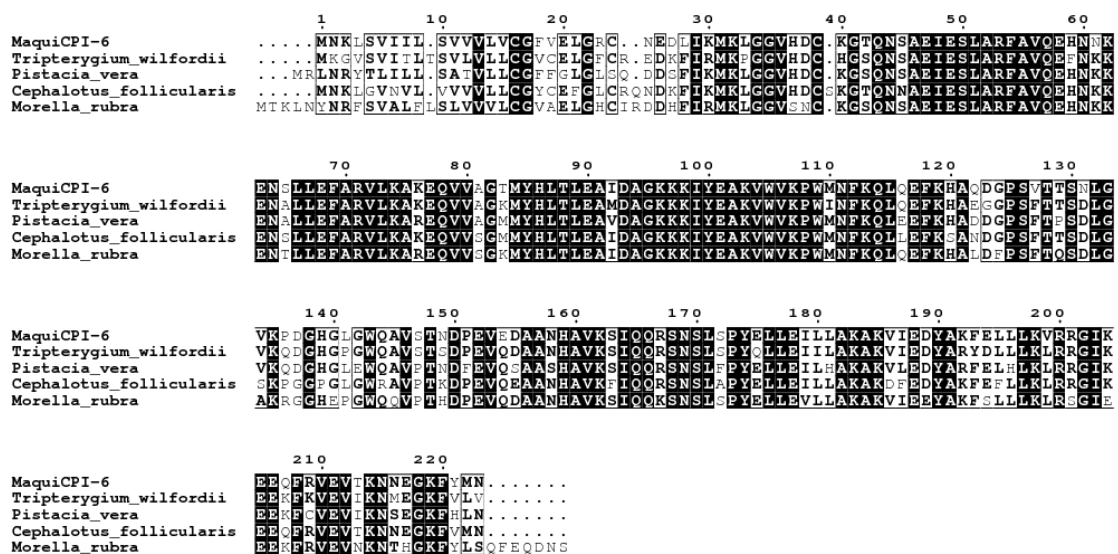


Figura 20: Alinhamento entre a MaquiCPI-6 e cistatinas de outras espécies vegetais. Foram alinhadas cistatinas de *Tripterygium wilfordii* (XP_038719056.1), *Pistacia vera* (XP_031272072.1), *Cephalotus follicularis* (GAV86994.1), *Morella rubra* (KAB1222999.1) com a MaquiCPI-6, utilizando o programa Muscle.

4.2 ANÁLISE FILOGENÉTICA

A análise filogenética entre as cistatinas de maqui ocorreu em função do alinhamento gerado na Figura 14. Foi utilizado o método de verossimilhança com o modelo WAG+G(2) (WHELAN; GOLDMAN, 2001). A qualidade da construção foi verificada a partir da análise de *bootstrap* em 1000 ciclos. O processo gerou a árvore não enraizada apresentada na Figura 21.

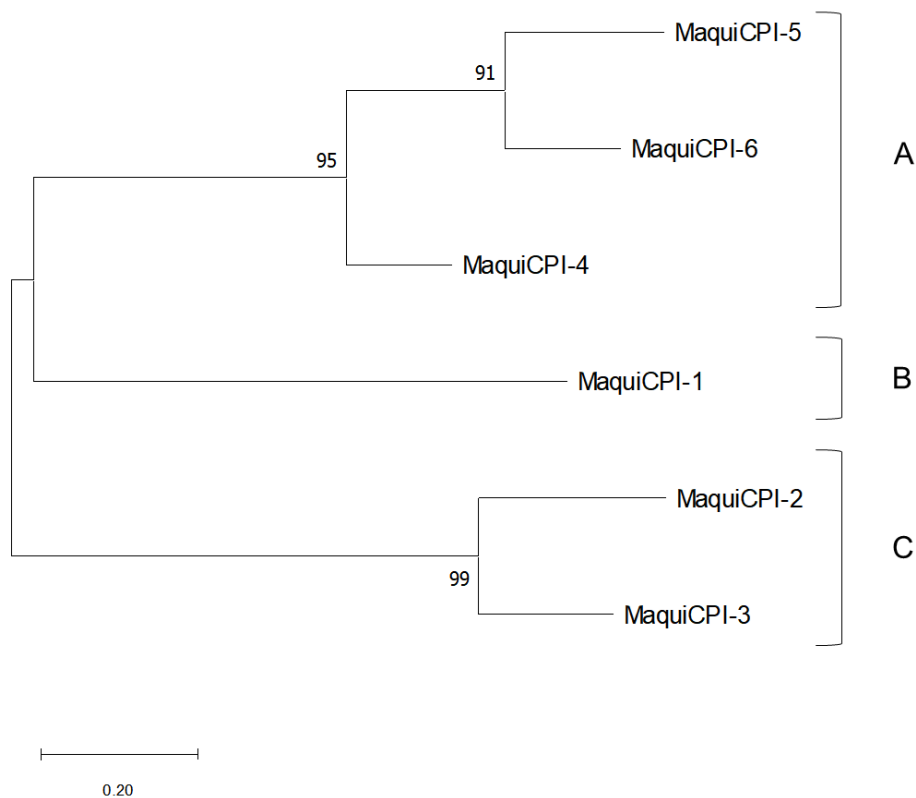


Figura 21: Árvore filogenética das maquicistatinas. A inferência filogenética foi realizada pelo método da máxima verossimilhança utilizando o modelo de WAG (WHELAN; GOLDMAN, 2001) com distribuição gamma de 2 categorias discretas. Os valores de *bootstrap* estão indicados nos nós como porcentagem. Os *clusters* estão indicados com as letras A, B e C.

As sequências de aminoácidos foram agrupadas em 3 grupos, assim como foi observado para citrus, nabo, arroz e cevada (CHRISTOFF; MARGIS, 2014; HONG *et al.*, 2008; MARTINEZ *et al.*, 2009; SCHNEIDER *et al.*, 2020). O grupo A é composto pela MaquiCPI-4 e um clado interno, que abriga a MaquiCPI-5 e MaquiCPI-6, que são estendidas. Elas apresentam um motivo da alfa hélice bastante conservado: LARFAV-[DEQ]-EHN. A MaquiCPI-1 é o único exemplo do grupo B, e é caracterizada por uma extensão de oito aminoácidos na posição 59 do alinhamento (Figura 14). Finalmente, o grupo C contempla as MaquiCPI-2 e MaquiCPI-3 que não possuem qualquer extensão e possuem alfa hélice com padrão IGEFAVD-[EA]-YN.

Para fins de comparação, uma árvore filogenética (Figura 22) foi construída entre as cistatinas de maqui e de arroz, espécie cujas cistatinas são bem descritas. Foi possível observar que não houve clusterização que separasse as sequências de maqui e de arroz. Na verdade, são formados os mesmos três grupos observados na árvore exclusivamente das maquicistatinas, apresentando as mesmas características, quando se refere a presença de extensões.

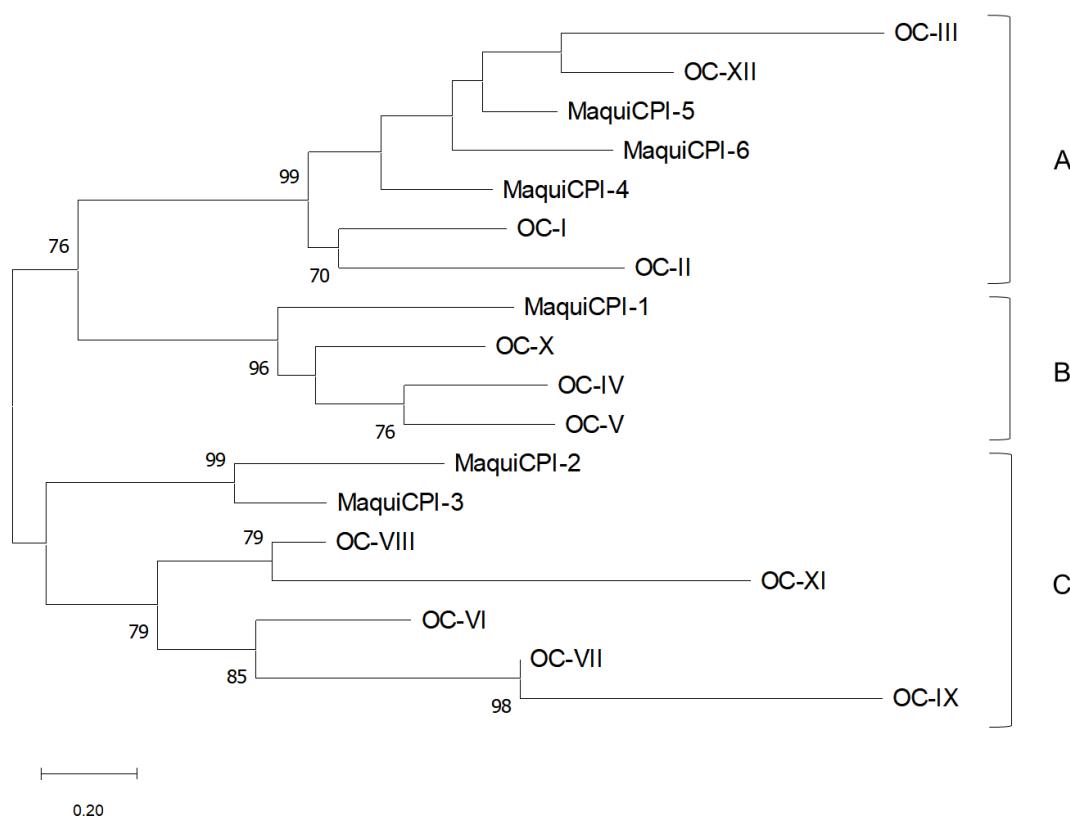


Figura 22: Árvore filogenética das cistatinas de maqui e de arroz (*Oriza sativa* spp *japonica*). A inferência filogenética foi realizada pelo método da máxima verossimilhança utilizando o modelo de WAG (WHELAN; GOLDMAN, 2001) com distribuição gamma de 2 categorias discretas. Os valores de *bootstrap* estão indicados nos nós como porcentagem. Os *clusters* estão indicados com as letras A, B e C.

4.3 ISOLAMENTO DAS ORFS E CONSTRUÇÃO DE VETORES DE EXPRESSÃO

Os primers sintetizados foram desenhados para amplificar as sequências de fitocistatinas excluindo as regiões de peptídeo sinal, quando presentes. Isso foi realizado pois, em estudos anteriores de nosso grupo de pesquisa, observou-se menor expressão e solubilidade das proteínas contendo o peptídeo sinal (GIANOTTI *et al.*, 2006). As ORFs de cada uma das cistatinas foram amplificadas por PCR e observadas em gel de agarose 1%, apresentando assim, o tamanho esperado para cada uma delas, considerando as bases adicionadas pelos primers: MaquiCPI-1 (345 pb), MaquiCPI-2 (306 pb), MaquiCPI-3 (303 pb), MaquiCPI-4 (321 pb) e MaquiCPI-5 (630 pb) (Figura 23). Não foi possível amplificar a sequência referente à

MaquiCPI-6, mesmo após inúmeras tentativas e variando as condições de reação. Como já havíamos amplificado a MaquiCPI-5, bastante similar, decidimos prosseguir os estudos apenas com ela.

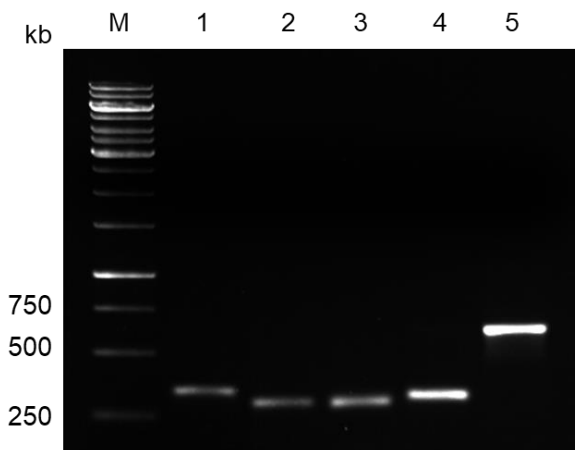


Figura 23: Amplificação das ORFs das maquicistatinas. (M) Marcador de peso molecular (GeneRuler 1 kb Thermoscientific), (1) MaquiCPI-1 - 345 pb, (2) MaquiCPI-2 - 306 pb, (3) MaquiCPI-3 - 303 pb, (4) MaquiCPI-4 - 321 pb e (5) MaquiCPI-5 - 630 pb.

Os produtos de PCR foram ligados primeiramente ao plasmídeo de propagação pTZ57R/T. A ligação nesse vetor é facilitada pois este possui uma base de timina em ambas extremidades 3' e os *amplicons* possuem um nucleotídeo adenina livre 3' devido a atividade da Taq polimerase. O sucesso da clonagem foi confirmado pela PCR de colônia e pela clivagem dos plasmídeos. Cada clone contendo os fragmentos de interesse foi clivado com as enzimas apropriadas para transferência ao vetor de expressão, pET28a. A clonagem neste vetor foi confirmada por PCR de colônia e restrição (Figura 24). A escolha desse vetor ocorreu devido ao sucesso na expressão e purificação de cistatinas de cana-de-açúcar pelo nosso grupo de pesquisa (GIANOTTI *et al.*, 2006; SOARES-COSTA *et al.*, 2002). Além disso, a utilização desse plasmídeo permite a fácil produção de proteínas a partir da indução por IPTG.

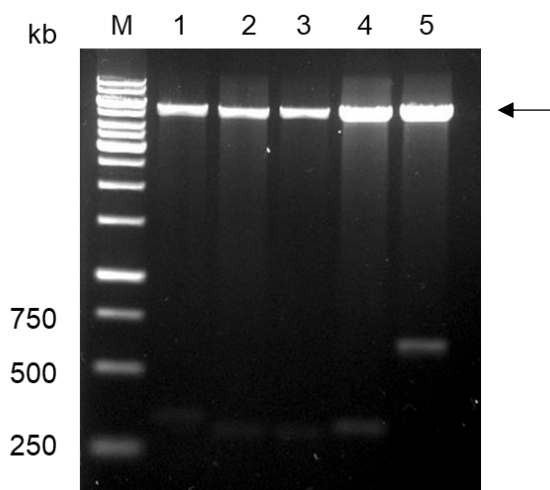


Figura 24: Confirmação da clonagem pela clivagem do plasmídeo pET28a+inserto. (M) Marcador de peso molecular (GeneRuler 1 kb Thermoscientific), (1) MaquiCPI-1 - 327 pb, (2) MaquiCPI-2 - 288 pb, (3) MaquiCPI-3 - 285 pb, (4) MaquiCPI-4 - 303 pb e (5) MaquiCPI-5 - 612 pb. O plasmídeo pET28-a (5369 pb) linearizado está indicado pela seta.

4.4 EXPRESSÃO E PURIFICAÇÃO DAS CISTATINAS RECOMBINANTES

As ORFs correspondentes as cistatinas foram clonadas em pET-28a. Esse plasmídeo permite a clonagem direcional do fragmento em fase com a sequência codificante de seis histidinas (his-tag) nas regiões N e C terminal. A sequência de histidina possibilita a purificação por cromatografia de afinidade e pode ser removida devido a um sítio de clivagem de trombina. Desse modo, as cistatinas foram expressas em fusão com a his-tag na região N-terminal. A Tabela 7 apresenta algumas das características das proteínas contendo a his-tag N-terminal, obtidas *in silico* como o peso molecular, pI e solubilidade em *E. coli*, comparada à proteína sem essa modificação. O programa SoluProt (HON *et al.*, 2021) define as proteínas solúveis nesse sistema de expressão aquelas com coeficiente superior a 0,5.

Tabela 7: Características das MaquiCPIs com ^(a) e sem ^(b) a his-tag N-terminal.

		Número de aminoácidos	Massa Molecular Teórica	pI	Solubilidade
MaquiCPI-1	(a)	130	14,45 kDa	9,10	0,675
	(b)	109	12,15 kDa	8,27	0,661
MaquiCPI-2	(a)	117	13,18 kDa	9,19	0,644
	(b)	96	10,89 kDa	8,97	0,505
MaquiCPI-3	(a)	116	12,67 kDa	7,11	0,650
	(b)	95	10,37 kDa	5,80	0,530
MaquiCPI-4	(a)	122	13,36 kDa	7,11	0,883
	(b)	101	11,06 kDa	5,83	0,917
MaquiCPI-5	(a)	225	25,06 kDa	8,41	0,887
	(b)	204	22,76 kDa	6,06	0,914

A análise por SDS-PAGE 15% (Figura 25) das amostras coletadas permitiu a avaliação do perfil de expressão proteica gradual a partir do início até a quarta hora de indução, bem como, das frações insolúvel, solúvel e purificada.

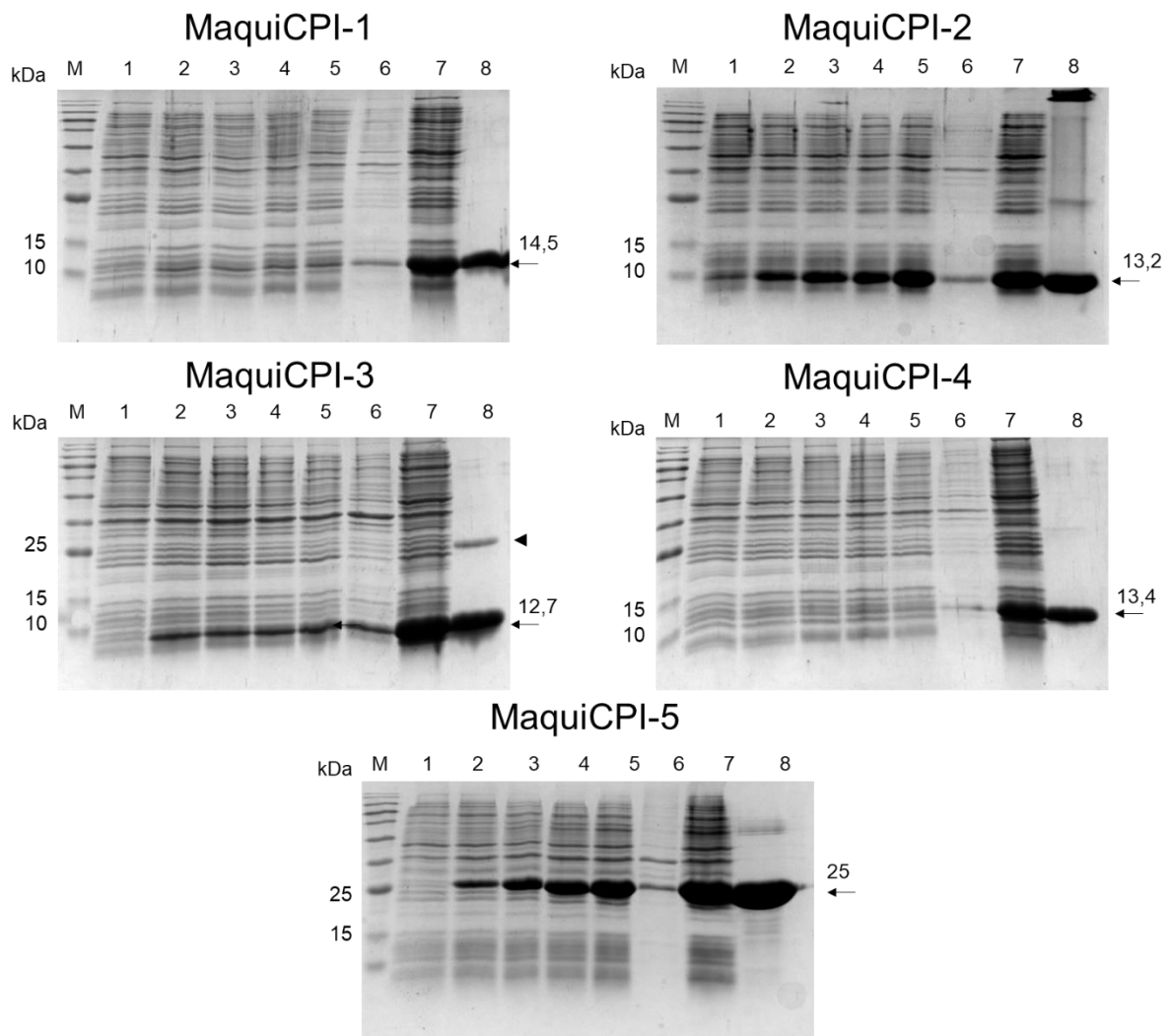


Figura 25: Expressão recombinante e purificação das maquicistatinas. Análise em SDS-PAGE 15% das amostras coletadas antes da indução (1), após 1 hora (2), 2 horas (3), 3 horas (4), 4 horas (5), das frações insolúvel (6) e solúvel (7) e da purificação (8). M: BlueClassic Prestained Protein Marker - Jena Bioscience. As setas apontam as proteínas expressas e a ponta de seta a provável formação de dímero na MaquiCPI-3.

As frações solúveis e insolúveis, resultantes da lise celular das bactérias, demonstram que, em todos os casos, as proteínas recombinantes apresentavam-se bastante representadas na fração solúvel (ainda que quase sempre, parte fique na fração insolúvel). As frações solúveis foram então aplicadas em colunas contendo resina de níquel para a purificação das cistatinas por cromatografia de afinidade. A eluição foi realizada com tampão de lise contendo concentrações crescentes de imidazol. Todas as cistatinas apresentaram eluição relevante da coluna a partir de 100 mM até 250 mM de imidazol.

O rendimento obtido para todos os casos foi bastante satisfatório (Tabela 8), pois as proteínas de interesse foram obtidas em concentrações e pureza suficientes para a realização dos ensaios experimentais. A diferença entre a concentração de proteínas obtidas após a purificação pode estar associada a diversos fatores, além da própria produtividade intrínseca de cada proteína em *E. coli*, que pode variar. A solubilidade pode afetar a quantidade disponível para o processo de purificação. De modo geral, as fitocistatinas possuem um núcleo hidrofóbico compacto formado entre a α -hélice e a folha- β formada pelas fitas- β antiparalelas. A presença de qualquer resíduo de aminoácido menos favorável, pode desestabilizar a interação entre a α -hélice e a folha- β , reduzindo a compactação e expondo alguns resíduos hidrofóbicos que podem levar a agregação e redução da solubilidade. A identificação desses resíduos pode ser realizada por modelagem molecular (VALADARES *et al.*, 2010). Outro fator é a capacidade de retenção na coluna de níquel, que pode afetar o rendimento final da proteína. Além disso, a natureza dos aminoácidos pode levar a diferenças estruturais que tornam a posição da sequência de histidinas mais propícias à ligação ao metal.

Tabela 8: Rendimento da produção das diferentes maquicistatinas.

	Rendimento por litro de cultura (mg)
MaquiCPI-1	17,54
MaquiCPI-2	23,61
MaquiCPI-3	84,14
MaquiCPI-4	18,46
MaquiCPI-5	44,64

Outra questão que está associada à produtividade durante a expressão de proteínas recombinantes é a natureza dos aminoácidos da região N-terminal (VERMA *et al.*, 2019). No entanto, todas as maquicistatinas foram expressas em fusão, no N-terminal, a 21 aminoácidos,

derivados do pET28a, incluindo a his-tag. Dessa forma, os resultados mostram que a influência é dos aminoácidos internos a cada uma das cistatinas e/ou seus arranjos estruturais.

Ao observar o perfil de expressão durante as 4 horas de indução da maquiCPI-4 comparando-a com a concentração por litro de cultura, nota-se um alto rendimento. A presença de bandas referentes à proteína nas canaletas de 2 a 5 da Figura 25 (D) é bastante sutil enquanto a canaleta 7, referente a sua fração solúvel, apresenta uma alta quantidade da proteína de interesse. Esse fenômeno deve estar associado ao seu alto índice de solubilidade (Tabela 7), resultante da sua composição de aminoácidos, conforme citado anteriormente.

Dentre as cistatinas de maqui, em termos de rendimento, se destaca a MaquiCPI-3, que rendeu mais de 80 mg por litro de cultura. Embora existam divergências nas metodologias empregadas para a produção, como meio de cultura e concentração de IPTG, pode-se observar que a quantidade produzida por litro de cultura é superior à maioria das cistatinas de plantas produzidas em *E. coli* descritas na literatura (Tabela 9). Observa-se também que há maior rendimento comparado a cistatinas humanas de interesse biotecnológico, como a estefina B e a cistatina C, evidenciando que possam ser substituídas por cistatinas vegetais mais rentáveis. Quando comparadas às cistatinas de cana-de-açúcar, tradicionalmente estudadas no Laboratório de Biologia Molecular da UFSCar, as oriundas de *Aristotelia chilensis* apresentam uma vantagem relevante quando se trata de quantidade e da ausência de precipitação em altas concentrações. A CaneCPI-5 que rende até 16 mg/L (SHIBAO *et al.*, 2021), aproximadamente, e precipita em concentrações acima de 1 mg/mL, tem exigido bastante tempo, mão de obra e reagentes, diante das demandas dos trabalhos em colaboração para uso da proteína. Portanto, caso apresentem propriedades similares, as maquicistatinas podem ser consideradas uma alternativa atrativa às que têm sido utilizadas. Diante disso, o alto rendimento desses inibidores de peptidases revela um alto potencial, não só para as atividades do laboratório, mas também para a indústria biotecnológica, farmacêutica e agrícola.

Tabela 9: Rendimento por litro de cultura de cistatinas de diferentes espécies

Nomenclatura	Rendimento	Referência
Feijão-caupi (<i>Vigna unguiculata</i>) - VuCys1	10 mg	(MONTEIRO JÚNIOR <i>et al.</i> , 2017)
Feijão-caupi (<i>Vigna unguiculata</i>) - VuCys2	22 mg	(MONTEIRO JÚNIOR <i>et al.</i> , 2017)
Caruru (<i>Amaranthus hypochondriacus</i>) – AhCPI	266 mg	(VALDES-RODRIGUEZ <i>et al.</i> , 2010)
Seringueira (<i>Hevea brasiliensis</i>) - HbCPI	30 mg	(BANGRAK; CHOTIGEAT, 2011)
Figueira (<i>Ficus awkeotsang</i>) - FaCYS	1,2 mg	(SHYU <i>et al.</i> , 2011)
Abacaxi (<i>Ananas comosus</i>) - AcCYS	63,7 mg	(SHU <i>et al.</i> , 2004)
<i>Arabidopsis thaliana</i> - AtCYS1	3,6 mg	(BELENGHI <i>et al.</i> , 2003)
Laranjeira (<i>Citrus sinensis</i>) - CsinCPI-2	10 mg	(SCHNEIDER <i>et al.</i> , 2020)
Clementina (<i>Citrus lementina</i>) - CClemCPI-2	10 mg	(SCHNEIDER <i>et al.</i> , 2020)
Cana-de-açúcar (<i>Saccharum officinarum</i>) - CaneCPI-1	10 mg	(SOARES-COSTA <i>et al.</i> , 2002)
Cana-de-açúcar (<i>Saccharum officinarum</i>) - CaneCPI-5	16 mg	(SANTIAGO <i>et al.</i> , 2017)
Humano (<i>Homo sapiens</i>) – Estefina B	4 mg	(JERALA <i>et al.</i> , 1988)
Humano (<i>Homo sapiens</i>) – Cistatina C	15 mg	(SZYMAŃSKA <i>et al.</i> , 2009)

Ao observar a fração purificada da MaquiCPI-3, nota-se uma banda que possui aparentemente o dobro de sua massa molecular. Trata-se da formação de homodímeros, um fenômeno comum entre as cistatinas. Provavelmente, isso foi decorrente do processo de *domain swap*, o que foi demonstrado para a CaneCPI-1 por meio de ensaios de NMR (ressonância magnética nuclear) (Figura 26) (VALADARES *et al.*, 2013). Esse fenômeno causa mudanças conformacionais no primeiro *loop* inibitório. Cada monômero emprega a alfa-hélice e a fita-beta 2 do outro, mantendo a conformação de fitocistatina. Adicionalmente, os resíduos do primeiro *loop* inibitório apresentam conformação de fita-beta, tornando as fitas-beta 2 e 3 em uma única e longa fita (Figura 27) (MONTEIRO JÚNIOR *et al.*, 2017; TURK; STOKA; TURK, 2008). Como os homodímeros são menos ativos, *in vivo*, este fenômeno provavelmente

está relacionado à regulação do controle da inibição endógena de cisteíno peptidase (OHTSUBO *et al.*, 2007).

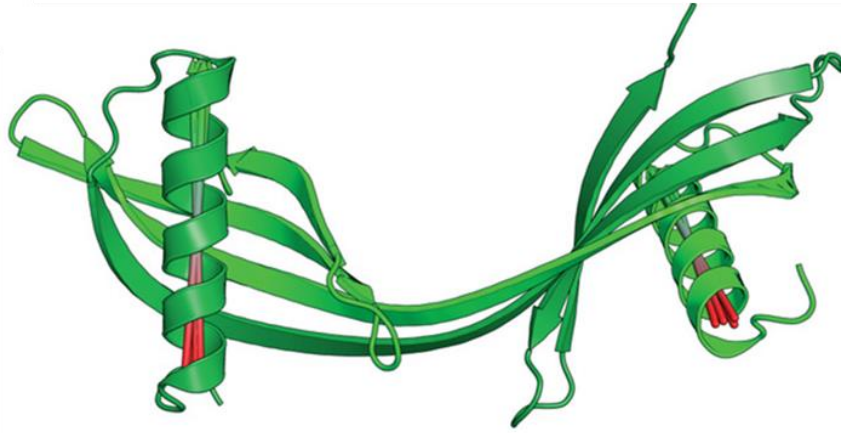


Figura 26: Representação da estrutura em dímero, por *domain swap*, da canecystatin-1. Estrutura obtida por cristalografia e NMR (PDB 3UL6) (VALADARES *et al.*, 2013).

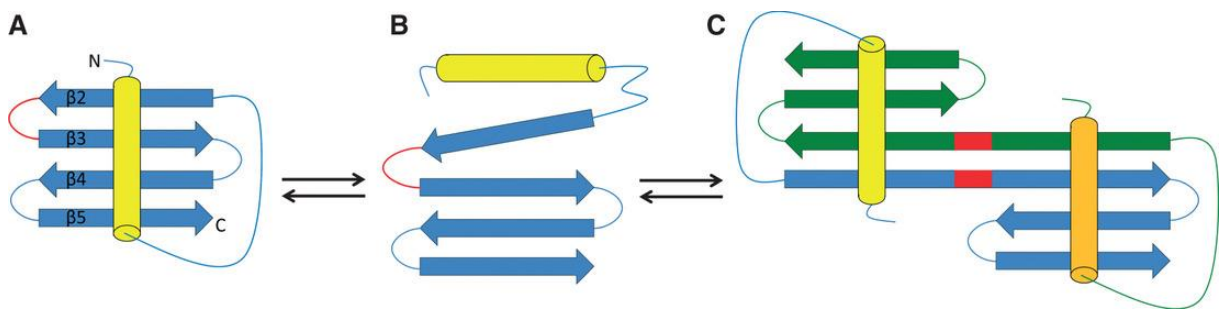


Figura 27: Esquema simplificado da transição entre o monômero fechado para o dímero por *domain swap*. Os cilindros amarelo e laranja representam alfa-hélices, e as setas azuis e verdes representam as fitas betas. (A) Monômero fechado. (B) monômero parcialmente desestruturado. (C) Dímero por *domain swap* (VALADARES *et al.*, 2013).

4.5 ENSAIOS DE INIBIÇÃO DE CISTEÍNO PEPTIDASES

A atividade inibitória das maquicistatinas recombinantes contra as cisteíno peptidases foi medida em espectrofluorímetro Hitachi F2500, utilizando o substrato fluorogênico Z-Phe-Arg-AMC. As atividades residuais das enzimas papaína e as catepsinas L e B humanas, à medida que concentrações constantes de inibidor foram adicionadas, podem ser observadas nos gráficos das Figura 28, Figura 29 e Figura 30, associados ao K_i aparente.

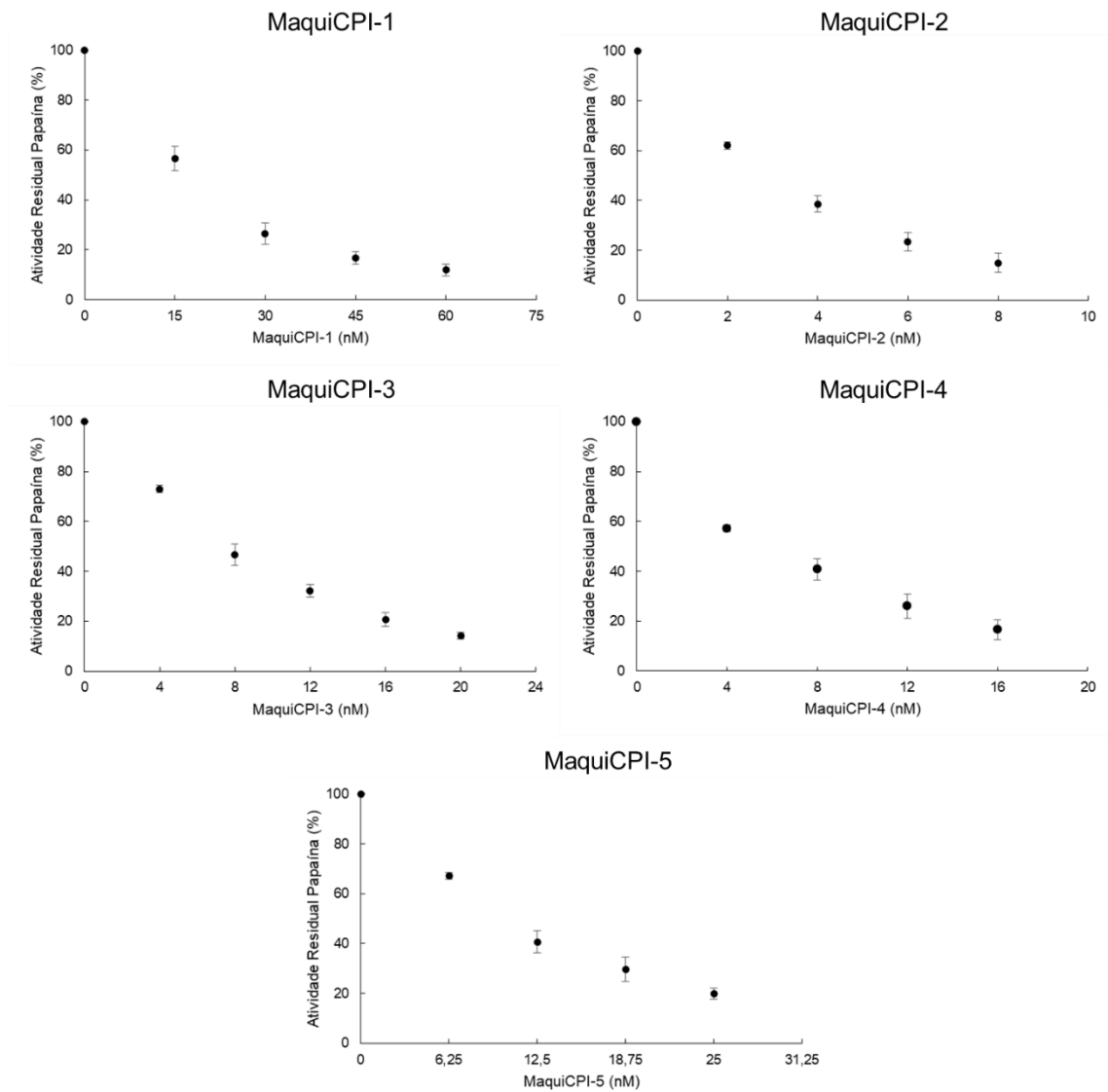


Figura 28: Inibição da Papaína (*Carica papaya*) pelas cistatinas recombinantes. A atividade inibitória das MaquiCPIs contra a enzima papaína foi medida em ensaio fluorimétrico utilizando o substrato Z-FR-MCA. Os gráficos indicam a curva de inibição de cada inibidor, obtida pelo cálculo da atividade residual da enzima à medida que o inibidor foi adicionado. Os valores estão expressos como a média \pm o desvio padrão de três experimentos independentes.

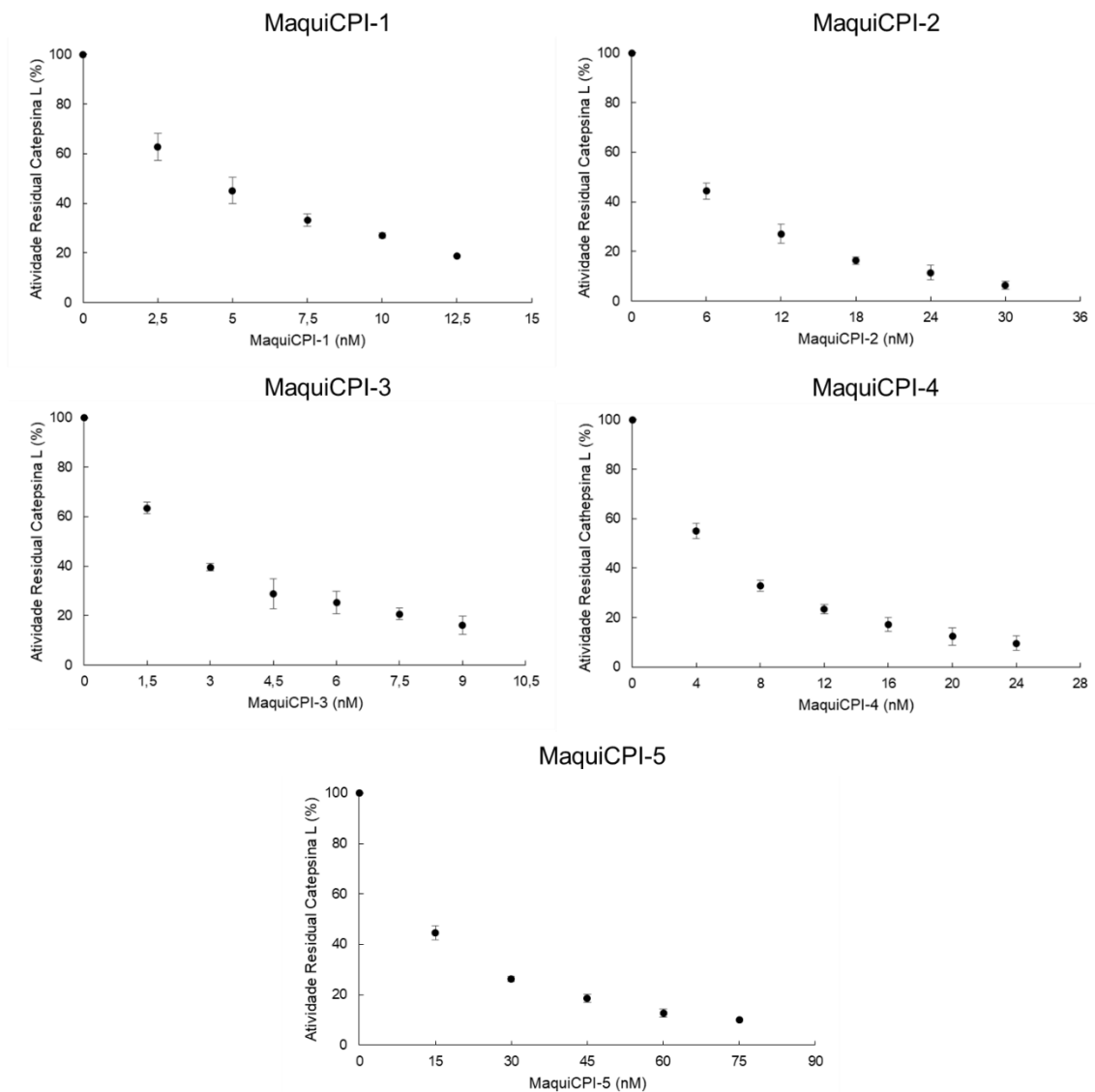


Figura 29: Inibição da Catepsina L humana pelas cistatinas recombinantes. A atividade inibitória das MaquiCPIs contra a enzima papaína foi medida em ensaio fluorimétrico utilizando o substrato Z-FR-MCA. Os gráficos indicam a curva de inibição de cada inibidor, obtida pelo cálculo da atividade residual da enzima à medida que o inibidor foi adicionado. Os valores estão expressos como a média \pm o desvio padrão de três experimentos independentes.

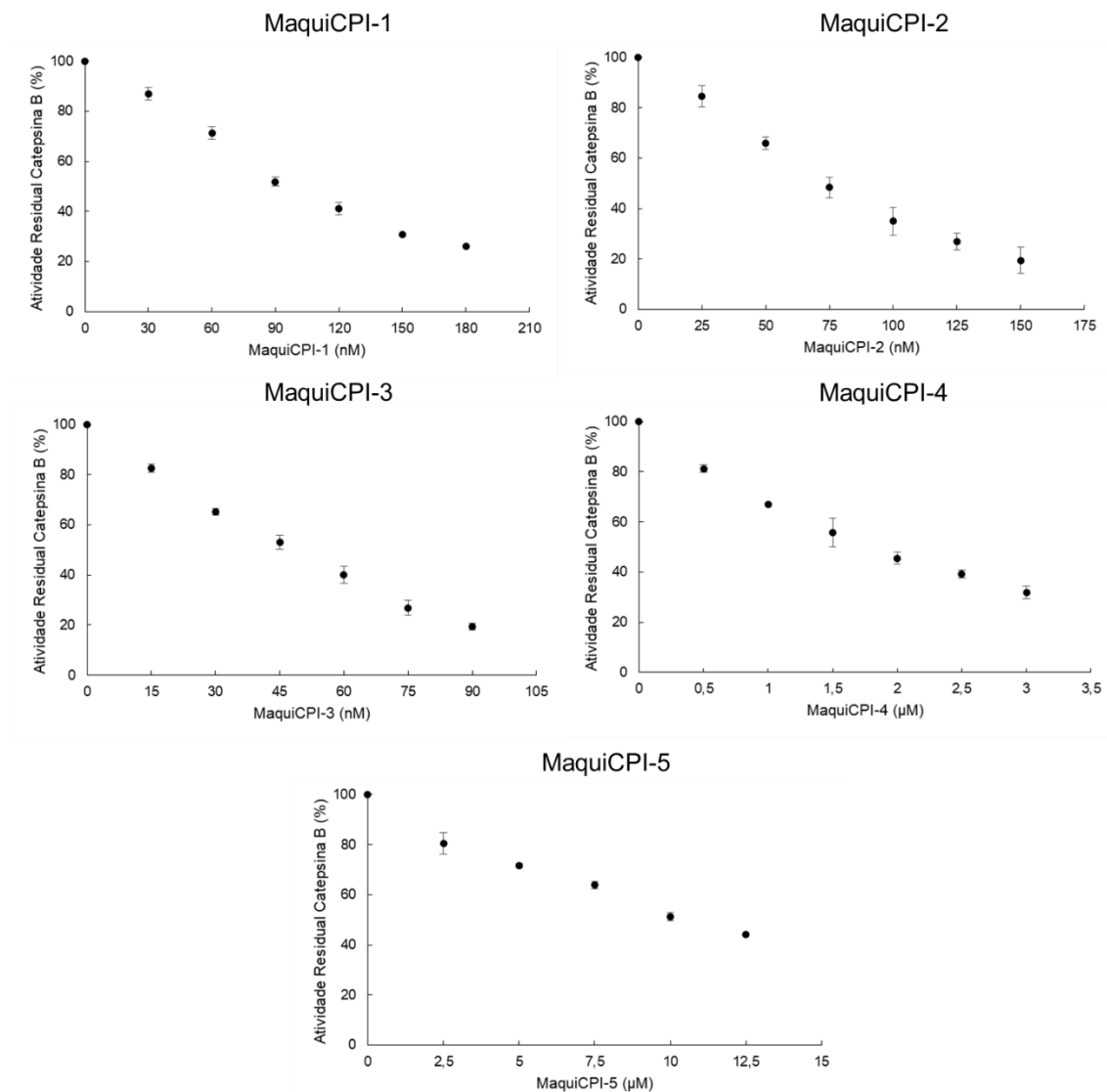


Figura 30: Inibição da Catepsina B humana pelas cistatinas recombinantes. A atividade inibitória das MaquiCPIs contra a enzima papaína foi medida em ensaio fluorimétrico utilizando o substrato Z-FR-MCA. Os gráficos indicam a curva de inibição de cada inibidor, obtida pelo cálculo da atividade residual da enzima à medida que o inibidor foi adicionado. Os valores estão expressos como a média \pm o desvio padrão de três experimentos independentes.

Os valores das constantes de inibição (Ki) estão dispostos na Tabela 10.

Tabela 10: Constantes de inibição (Ki) das maquicistatinas contra a papaína e as catepsinas B e L humanas.

	Constante de Inibição – Ki (nM)		
	Papaína	Catepsina L	Catepsina B
MaquiCPI-1	7,13±1,30	0,34±0,07	35,74±0,84
MaquiCPI-2	1,42±0,39	0,33±0,07	20,97±3,78
MaquiCPI-3	3,29±0,58	0,38±0,08	21,94±3,01
MaquiCPI-4	2,99±0,34	0,57±0,11	876,70±53,99
MaquiCPI-5	5,05±0,82	1,25±0,25	5470±630

As cistatinas MaquiCPI-1, MaquiCPI-1, MaquiCPI-2, MaquiCPI-3, MaquiCPI-4 e MaquiCPI-5 foram capazes de inibir de maneira eficiente a papaína, peptidase modelo para estudos de inibição de cisteíno peptidases, com valores de Ki de 7,13 nM, 1,42 nM, 3,29 nM, 2,99 nM e 5,05 nM, respectivamente. A sua capacidade inibitória indica que as maquicistatinas recombinantes foram produzidas em sua conformação correta e comprovam que são inibidores de cisteíno peptidases.

Quando se trata da inibição da catepsina B, nota-se que as cistatinas MaquiCPIs 1, 2 e 3, são capazes de inibi-la na ordem de nanomolar, com Ki de 35,74 nM, 20,97 nM e 21,94 nM, respectivamente. No entanto, as MaquiCPIs 4 e 5 têm um potencial inibitório bem baixo, com Ki na ordem de micromolar (0,876 μ M e 5,47 μ M). O fato de algumas cistatinas serem incapazes de inibir a catepsina B, está associado ao bloqueio da fenda do sítio ativo da peptidase por seu *loop* de oclusão (MUSIL *et al.*, 1991). Isso, portanto, reflete na capacidade de interação das cistatinas e, conseqüentemente, valores altos de Ki para as MaquiCPIs 4 e 5.

Quando ocorre a inibição da catepsina B pelas cistatinas, é necessário um processo de duas etapas, no qual há uma mudança conformacional em que o *loop* de oclusão é deslocado pelo inibidor. Isso desencadeia a desobstrução do sítio ativo, permitindo a inibição (NYCANDER *et al.*, 1998; PAVLOVA *et al.*, 2000). Esse mecanismo está associado à região N-terminal das fitocistatinas (BJÖRK *et al.*, 1994; HALL *et al.*, 1998). No entanto, esse efeito não se limita a essa região, o que foi comprovado com experimentos com as cistatinas de cana-

de-açúcar. Ao se adicionar a região N-terminal da CaneCPI-4, potente inibidor de catepsina B, às CaneCPIs 2 e 3, verificou-se que estas não se tornaram capazes de inibir a peptidase (GIANOTTI *et al.*, 2008).

A catepsina L, por sua vez, foi inibida com sucesso por todas as cistatinas testadas e, com exceção da MaquiCPI-5, apresentam valores de K_i inferiores a 1 nM. Isso reflete em seu potencial biotecnológico associado ao controle de patologias resultantes da quantidade ou função desregulada dessa peptidase. Com a administração de cistatinas nesses casos problemas neurológicos como o mal de Parkinson (LI *et al.*, 2011), urológicos como a proteinúria (CAO *et al.*, 2017) e até mesmo cânceres (SUDHAN; SIEMANN, 2015), poderiam ser controlados. Atualmente, a inibição da catepsina L vem ganhando destaque como estratégia para a inibição do ciclo do SARSCoV-2, pois ela é fundamental para uma das vias de entrada do vírus na célula do hospedeiro (GOMES *et al.*, 2020).

O trabalho mostrou, pela primeira vez, a identificação de cistatinas de uma planta nativa do Chile, ainda pouco estudada, que possui grande apelo comercial e científico. As proteínas foram produzidas de maneira recombinante em quantidades bastante satisfatórias e de maneira solúvel, além de apresentarem poder inibitório expressivo contra cisteíno peptidases. Isso revela um grande potencial biotecnológico e comercial das proteínas de estudo, abrangendo áreas como a medicina, odontologia e agricultura.

5 CONCLUSÕES

- O transcriptoma de maqui, realizado por meio de sequenciamento de nova geração utilizando RNA de plantas cultivadas em TIBs, permitiu a identificação de seis diferentes cistatinas, das quais cinco tiveram a sequência amplificada a partir do cDNA.
- A expressão das cistatinas, combinando a utilização do vetor pET28-a e a cepa de *E. coli* Rosetta (DE3) se mostrou bastante eficiente, possibilitando a obtenção de grande quantidade das proteínas nas frações solúveis, com destaque para a maquiCPI-3, o que permitiu a sua purificação resultando em proteína suficiente para a execução dos ensaios de atividade.
- As proteínas foram expressas em suas conformações corretas, o que foi confirmado pelos ensaios de inibição da papaína e outras cisteíno peptidases.
- Todas as cistatinas foram capazes de inibir a catepsina L com Ki na faixa de nanomolar. No entanto, somente as MaquiCPIs 1, 2 e 3 foram capazes de inibir a catepsina B nessa faixa.
- A inibição das catepsinas lisossomais refletem o potencial das cistatinas de maqui em aplicações na área da saúde, a partir do seu direcionamento a doenças associadas a superexpressão dessas proteases.

6 PERSPECTIVAS E TRABALHOS SENDO REALIZADOS

- As cistatinas já foram anteriormente descritas como sendo capazes de inibir a replicação de vírus da família Coronaviridae em cultura celular (COLLINS; GRUBB, 1991, 1998). As maquicistatinas mostraram-se eficientes em inibir as catepsinas lisossomais humanas B e L, que estão associadas à entrada do SARS-CoV-2 no ambiente intracelular. Diante disso, em parceria com a Profa. Dra. Carolina Borsoi Moraes do ICB (USP), ensaios celulares de inibição da proliferação viral com as MaquiCPIs 3 e 4 vêm sendo realizados, com resultados promissores.
- Cistatinas de cana-de-açúcar têm sido amplamente utilizadas em colaborações com foco em tratamentos odontológicos, porém o seu rendimento é relativamente baixo, o que dificulta a obtenção de proteínas em quantidade suficiente para atender a demanda. Por essa razão, a MaquiCPI-3, que é produzida em quantidade bastante superior às de cana-de-açúcar, e sem que haja precipitação em altas concentrações, vem sendo estudada quanto ao seu potencial de se ligar ao esmalte dentário, com resultados promissores.
- Recentemente, nosso grupo de pesquisa, em colaboração com o grupo de pesquisa da Profa. Dra. Fernanda Araújo da Universidade Federal de Uberlândia, verificou que uma cistatina de cana-de-açúcar (CaneCPI-5) tem atividade anti-inflamatória e angiogênica, além de aumentar a deposição de colágeno. Nesse sentido, as maquicistatinas também serão testadas como agentes anti-inflamatórios e em reparo de feridas, usando modelo de implante subcutâneo.
- As amostras de RNA utilizadas no trabalho foram oriundas de um experimento que avaliou o efeito do ácido abscísico no cultivo de plântulas de maqui em biorreatores de imersão temporária. Foram coletadas amostras com e sem o fitorregulador no início do experimento e após 24 e 48 horas, que tiveram seu RNA total extraído e sequenciado. Nesse cenário, a partir do sequenciamento será avaliado o perfil de expressão das maquicistatinas no decorrer do tempo sob a influência do ABA.

REFERÊNCIAS

- ABRAHAMSON, M.; RITONJA, A.; BROWN, M. A.; GRUBB, A.; MACHLEIDT, W.; BARRETT, A. J. Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. **Journal of Biological Chemistry**, v. 262, n. 20, p. 9688–9694, 1987. Disponível em: [https://doi.org/10.1016/s0021-9258\(18\)47989-6](https://doi.org/10.1016/s0021-9258(18)47989-6)
- ALMAGRO ARMENTEROS, J. J.; TSIRIGOS, K. D.; SØNDERBY, C. K.; PETERSEN, T. N.; WINTHER, O.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. **Nature Biotechnology**, v. 37, n. 4, p. 420–423, 2019. Disponível em: <https://doi.org/10.1038/s41587-019-0036-z>
- ALVAREZ-FERNANDEZ, M.; BARRETT, A. J.; GERHARTZ, B.; DANDO, P. M.; NI, J.; ABRAHAMSON, M. Inhibition of Mammalian Legumain by Some Cystatins Is Due to a Novel Second Reactive Site. **Journal of Biological Chemistry**, v. 274, n. 27, p. 19195–19203, 1999. Disponível em: <https://doi.org/10.1074/jbc.274.27.19195>
- ARAI, S.; ABE, K. Cystatin-based control of insects, with special reference to oryzacystatin. *In*: MICHAUD, D. (org.). **Recombinant Protease Inhibitors in Plants**. Georgetown: Landes Bioscience, 2000. p. 27–40.
- ARAUJO, T. T. *et al.* A sugarcane cystatin (CaneCPI-5) alters microcosm biofilm formation and reduces dental caries. **Biofouling**, v. 37, n. 1, p. 109–116, 2021. Disponível em: <https://doi.org/10.1080/08927014.2021.1881065>
- BANGRAK, P.; CHOTIGEAT, W. Molecular cloning and biochemical characterization of a novel cystatin from Hevea rubber latex. **Plant Physiology and Biochemistry**, v. 49, n. 3, p. 244–250, 2011. Disponível em: <https://doi.org/10.1016/j.plaphy.2010.12.007>
- BANNAI, H.; TAMADA, Y.; MARUYAMA, O.; NAKAI, K.; MIYANO, S. Extensive feature detection of N-terminal protein sorting signals. **Bioinformatics**, v. 18, n. 2, p. 298–305, 2002. Disponível em: <https://doi.org/10.1093/bioinformatics/18.2.298>
- BARRETT, A. J. The cystatins: a new class of peptidase inhibitors. **Trends in Biochemical Sciences**, v. 12, p. 193–196, 1987. Disponível em: [https://doi.org/10.1016/0968-0004\(87\)90092-2](https://doi.org/10.1016/0968-0004(87)90092-2)
- BARRETT, A. J. Proteases. **Current Protocols in Protein Science**, v. 21, n. 1, 2000.

Disponível em: <https://doi.org/10.1002/0471140864.ps2101s21>

BARRETT, A. J. Proteolytic enzymes: nomenclature and classification. *In*: BOND, J. S.; BEYNON, R. J. (org.). **Proteolytic enzymes: A practical approach**. 2. ed. New York: Oxford, 2001. p. 1–21.

BARRETT, A. J.; MCDONALD, J. K. Nomenclature: protease, proteinase and peptidase. **Biochemical Journal**, v. 237, n. 3, p. 935–935, 1986. Disponível em: <https://doi.org/10.1042/bj2370935>

BARRETT, A. J.; RAWLINGS, N. D.; DAVIES, M. E.; MACHLEIDT, W.; SALVESEN, G.; TURK, V. Cysteine protease inhibitors of the cystatin superfamily. *In*: BARRETT, A. J.; SALVESEN, G. (org.). **Proteinase Inhibitors**. Amsterdam: Elsevier, 1986. p. 529–569.

BARRETT, A. J.; RAWLINGS, N.; WOESSNER, J. **Handbook of Proteolytic Enzymes**. 3. ed. Cambridge: Academic Press, 2012.

BELENGHI, B.; ACCONCIA, F.; TROVATO, M.; PERAZZOLLI, M.; BOCEDI, A.; POLITICELLI, F.; ASCENZI, P.; DELLEDONNE, M. AtCYS1, a cystatin from *Arabidopsis thaliana*, suppresses hypersensitive cell death. **European Journal of Biochemistry**, v. 270, n. 12, p. 2593–2604, 2003. Disponível em: <https://doi.org/10.1046/j.1432-1033.2003.03630.x>

BENCHABANE, M.; SCHLÜTER, U.; VORSTER, J.; GOULET, M. C.; MICHAUD, D. Plant cystatins. **Biochimie**, v. 92, n. 11, p. 1657–1666, 2010. Disponível em: <https://doi.org/10.1016/j.biochi.2010.06.006>

BERGER, A.; SCHECHTER, I. Mapping the active site of papain with the aid of peptide substrates and inhibitors. **Philosophical Transactions of the Royal Society of London. B, Biological Sciences**, v. 257, n. 813, p. 249–264, 1970. Disponível em: <https://doi.org/10.1098/rstb.1970.0024>

BJÖRK, I.; POL, E.; RAUB-SEGALL, E.; ABRAHAMSON, M.; ROWAN, A. D.; MORT, J. S. Differential changes in the association and dissociation rate constants for binding of cystatins to target proteinases occurring on N-terminal truncation of the inhibitors indicate that the interaction mechanism varies with different enzymes. **Biochemical Journal**, v. 299, n. 1, p. 219–225, 1994. Disponível em: <https://doi.org/10.1042/bj2990219>

BODE, W.; ENGH, R.; MUSIL, D.; THIELE, U.; HUBER, R.; KARSHIKOV, A.; BRZIN, J.; KOS, J.; TURK, V. The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its

possible mode of interaction with cysteine proteinases. **The EMBO Journal**, v. 7, n. 8, p. 2593–2599, 1988.

BRAUCH, J. E. Underutilized Fruits and Vegetables as Potential Novel Pigment Sources. *In*: CARLE, R.; SCHWEIGGERT, R. M. (org.). **Handbook on Natural Pigments in Food and Beverages**. Stuttgart: Woodhead Publishing, 2016. p. 305–335. *E-book*. Disponível em: <https://doi.org/10.1016/B978-0-08-100371-8.00015-4>

BRAUCH, J. E.; BUCHWEITZ, M.; SCHWEIGGERT, R. M.; CARLE, R. Detailed analyses of fresh and dried maqui (*Aristotelia chilensis* (Mol.) Stuntz) berries and juice. **Food Chemistry**, v. 190, p. 308–316, 2016. Disponível em: <https://doi.org/10.1016/j.foodchem.2015.05.097>

CAO, Y.; LIU, X.; LI, Y.; LU, Y.; ZHONG, H.; JIANG, W.; CHEN, A. F.; BILLIAR, T. R.; YUAN, H.; CAI, J. Cathepsin L activity correlates with proteinuria in chronic kidney disease in humans. **International Urology and Nephrology**, v. 49, n. 8, p. 1409–1417, 2017. Disponível em: <https://doi.org/10.1007/s11255-017-1626-7>

CAVALLO-MEDVED, D.; MOIN, K.; SLOANE, B. Cathepsin B: Basis Sequence: Mouse. **The AFCS-nature molecule pages**, p. A000508, 2011. Disponível em: <https://doi.org/https://www.ncbi.nlm.nih.gov/pubmed/28781583>

CÉSPEDES, C. L.; ALARCON, J.; AVILA, J. G.; NIETO, A. Anti-inflammatory Activity of *Aristotelia chilensis* Mol. (Stuntz) (Elaeocarpaceae) . **Boletín Latinoamericano y del Caribe de Plantas Medicinales y Aromáticas**, v. 9, n. 2, p. 127–135, 2010.

CÉSPEDES, C. L.; EL-HAFIDI, M.; PAVON, N.; ALARCON, J. Antioxidant and cardioprotective activities of phenolic extracts from fruits of Chilean blackberry *Aristotelia chilensis* (Elaeocarpaceae), Maqui. **Food Chemistry**, v. 107, n. 2, p. 820–829, 2008. Disponível em: <https://doi.org/10.1016/j.foodchem.2007.08.092>

CHAN, S. J.; SAN SEGUNDO, B.; MCCORMICK, M. B.; STEINER, D. F. Nucleotide and predicted amino acid sequences of cloned human and mouse preprocathepsin B cDNAs. **Proceedings of the National Academy of Sciences**, v. 83, n. 20, p. 7721–7725, 1986. Disponível em: <https://doi.org/10.1073/pnas.83.20.7721>

CHRISTOFF, A. P.; MARGIS, R. The diversity of rice phytocystatins. **Molecular Genetics and Genomics**, v. 289, n. 6, p. 1321–1330, 2014. Disponível em: <https://doi.org/10.1007/s00438-014-0892-7>

COLLINS, A. R.; GRUBB, A. Inhibitory effects of recombinant human cystatin C on human coronaviruses. **Antimicrobial Agents and Chemotherapy**, v. 35, n. 11, p. 2444–2446, 1991. Disponível em: <https://doi.org/10.1128/AAC.35.11.2444>

COLLINS, A. R.; GRUBB, A. Cystatin D, a natural salivary cysteine protease inhibitor, inhibits coronavirus replication at its physiologic concentration. **Oral Microbiology and Immunology**, v. 13, n. 1, p. 59–61, 1998. Disponível em: <https://doi.org/10.1111/j.1399-302X.1998.tb00753.x>

COULOMBE, R.; GROCHULSKI, P.; SIVARAMAN, J.; MÉNARD, R.; MORT, J. S.; CYGLER, M. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. **The EMBO journal**, v. 15, n. 20, p. 5492–503, 1996. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/8896443>

DA COSTA FERNANDES, C.; RODRÍGUEZ, V. M. O.; SOARES-COSTA, A.; CIRELLI, J. A.; JUSTINO, D. M. N.; ROMA, B.; ZAMBUZZI, W. F.; FARIA, G. Cystatin-like protein of sweet orange (CsinCPI-2) modulates pre-osteoblast differentiation via β -Catenin involvement. **Journal of Materials Science: Materials in Medicine**, v. 32, n. 4, p. 33, 2021. Disponível em: <https://doi.org/10.1007/s10856-021-06504-y>

DANA, D.; PATHAK, S. K. A Review of Small Molecule Inhibitors and Functional Probes of Human Cathepsin L. **Molecules**, v. 25, n. 3, p. 698, 2020. Disponível em: <https://doi.org/10.3390/molecules25030698>

DILLARD, C. J.; GERMAN, J. B. Phytochemicals: nutraceuticals and human health. **Journal of the Science of Food and Agriculture**, v. 80, n. 12, p. 1744–1756, 2000. Disponível em: [https://doi.org/10.1002/1097-0010\(20000915\)80:12<1744::AID-JSFA725>3.0.CO;2-W](https://doi.org/10.1002/1097-0010(20000915)80:12<1744::AID-JSFA725>3.0.CO;2-W)

DRENTH, J.; JANSONIUS, J. N.; KOEKOEK, R.; SWEN, H. M.; WOLTHERS, B. G. Structure of Papain. **Nature**, v. 218, n. 5145, p. 929–932, 1968. Disponível em: <https://doi.org/10.1038/218929a0>

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p. 1792–1797, 2004. Disponível em: <https://doi.org/10.1093/nar/gkh340>

FAN, Y. S.; BYERS, M. G.; EDDY, R. L.; JOSEPH, L.; SUKHATME, V.; CHAN, S. J.; SHOWS, T. B. Cathepsin-L (CtsL) is located in the chromosome-9q21-q22 region-a related sequence is located on chromosome-10. **Cytogenetics and Cell Genetics**, v. 51, n. 1–4, p. 996,

1989.

FERREIRA, B. A.; TOYAMA, D.; HENRIQUE-SILVA, F.; ARAÚJO, F. de A. Recombinant sugarcane cystatin CaneCPI-5 down regulates inflammation and promotes angiogenesis and collagen deposition in a mouse subcutaneous sponge model. **International Immunopharmacology**, v. 96, 2021. Disponível em: <https://doi.org/10.1016/j.intimp.2021.107801>

FREIRE, M. C. L. C. *et al.* Non-Toxic Dimeric Peptides Derived from the Bothropstoxin-I Are Potent SARS-CoV-2 and Papain-like Protease Inhibitors. **Molecules**, v. 26, n. 16, p. 4896, 2021. Disponível em: <https://doi.org/10.3390/molecules26164896>

FUJISHIMA, A.; IMAI, Y.; NOMURA, T.; FUJISAWA, Y.; YAMAMOTO, Y.; SUGAWARA, T. The crystal structure of human cathepsin L complexed with E-64. **FEBS Letters**, v. 407, n. 1, p. 47–50, 1997. Disponível em: [https://doi.org/10.1016/S0014-5793\(97\)00216-0](https://doi.org/10.1016/S0014-5793(97)00216-0)

GASTEIGER, E. ExPASy: the proteomics server for in-depth protein knowledge and analysis. **Nucleic Acids Research**, v. 31, n. 13, p. 3784–3788, 2003. Disponível em: <https://doi.org/10.1093/nar/gkg563>

GIANOTTI, A.; RIOS, W. M.; SOARES-COSTA, A.; NOGAROTO, V.; CARMONA, A. K.; OLIVA, M. L. V.; ANDRADE, S. S.; HENRIQUE-SILVA, F. Recombinant expression, purification, and functional analysis of two novel cystatins from sugarcane (*Saccharum officinarum*). **Protein Expression and Purification**, v. 47, n. 2, p. 483–489, 2006. Disponível em: <https://doi.org/10.1016/j.pep.2005.10.026>

GIANOTTI, A.; SOMMER, C. A.; CARMONA, A. K.; HENRIQUE-SILVA, F. Inhibitory effect of the sugarcane cystatin CaneCPI-4 on cathepsins B and L and human breast cancer cell invasion. **Biological Chemistry**, v. 389, n. 4, p. 447–453, 2008. Disponível em: <https://doi.org/10.1515/BC.2008.035>

GOMES, C. P.; FERNANDES, D. E.; CASIMIRO, F.; DA MATA, G. F.; PASSOS, M. T.; VARELA, P.; MASTROIANNI-KIRSZTAJN, G.; PESQUERO, J. B. Cathepsin L in COVID-19: From Pharmacological Evidences to Genetics. **Frontiers in Cellular and Infection Microbiology**, v. 10, 2020. Disponível em: <https://doi.org/10.3389/fcimb.2020.589505>

GOULET, M.-C.; DALLAIRE, C.; VAILLANCOURT, L.-P.; KHALF, M.; BADRI, A. M.; PRERADOV, A.; DUCEPPE, M.-O.; GOULET, C.; CLOUTIER, C.; MICHAUD, D. Tailoring

the Specificity of a Plant Cystatin toward Herbivorous Insect Digestive Cysteine Proteases by Single Mutations at Positively Selected Amino Acid Sites. **Plant Physiology**, v. 146, n. 3, p. 1010–1019, 2008. Disponible em: <https://doi.org/10.1104/pp.108.115741>

HALL, A.; EKIEL, I.; MASON, R. W.; KASPRZYKOWSKI, F.; GRUBB, A.; ABRAHAMSON, M. Structural Basis for Different Inhibitory Specificities of Human Cystatins C and D. **Biochemistry**, v. 37, n. 12, p. 4071–4079, 1998. Disponible em: <https://doi.org/10.1021/bi971197j>

HASHIMOTO, Y.; KONDO, C.; KATUNUMA, N. An Active 32-kDa Cathepsin L Is Secreted Directly from HT 1080 Fibrosarcoma Cells and Not via Lysosomal Exocytosis. **PLOS ONE**, v. 10, n. 12, p. e0145067, 2015. Disponible em: <https://doi.org/10.1371/journal.pone.0145067>

HASNAIN, S.; HIRAMA, T.; HUBER, C. P.; MASON, P.; MORT, J. S. Characterization of cathepsin B specificity by site-directed mutagenesis. Importance of Glu245 in the S2-P2 specificity for arginine and its role in transition state stabilization. **The Journal of biological chemistry**, v. 268, n. 1, p. 235–40, 1993. Disponible em: <http://www.ncbi.nlm.nih.gov/pubmed/8093241>

HE, K.; VALCIC, S.; TIMMERMANN, B. N.; MONTENEGRO, G. Indole Alkaloids from *Aristolelia chilensis* (Mol.) Stuntz. **International Journal of Pharmacognosy**, v. 35, n. 3, p. 215–217, 1997. Disponible em: <https://doi.org/10.1076/phbi.35.3.215.13297>

HILLER, K.; GROTE, A.; SCHEER, M.; MUNCH, R.; JAHN, D. PrediSi: prediction of signal peptides and their cleavage positions. **Nucleic Acids Research**, v. 32, n. Web Server, p. W375–W379, 2004. Disponible em: <https://doi.org/10.1093/nar/gkh378>

HOFFMANN, A.; FARGA, C.; LASTRA, J.; VEGHAZI, E. **Plantas Medicinales de uso común en Chile**. Santiago: Ediciones Claudio Gay, 1992.

HON, J.; MARUSIAK, M.; MARTINEK, T.; KUNKA, A.; ZENDULKA, J.; BEDNAR, D.; DAMBORSKY, J. SoluProt: prediction of soluble protein expression in *Escherichia coli*. **Bioinformatics**, v. 37, n. 1, p. 23–28, 2021. Disponible em: <https://doi.org/10.1093/bioinformatics/btaa1102>

HONG, J.-K. *et al.* Identification and characterization of the phytocystatin family from *Brassica rapa*. **Journal of Plant Biotechnology**, v. 35, n. 4, p. 317–327, 2008. Disponible em: <https://doi.org/10.5010/JPB.2008.35.4.317>

HWANG, J. E.; HONG, J. K.; JE, J. H.; LEE, K. O.; KIM, D. Y.; LEE, S. Y.; LIM, C. O. Regulation of seed germination and seedling growth by an Arabidopsis phytocystatin isoform, AtCYS6. **Plant Cell Reports**, v. 28, n. 11, p. 1623–1632, 2009. Disponível em: <https://doi.org/10.1007/s00299-009-0762-7>

ILLY, C.; QURAIISHI, O.; WANG, J.; PURISIMA, E.; VERNET, T.; MORT, J. S. Role of the Occluding Loop in Cathepsin B Activity. **Journal of Biological Chemistry**, v. 272, n. 2, p. 1197–1202, 1997. Disponível em: <https://doi.org/10.1074/jbc.272.2.1197>

JERALA, R.; TRSTENJAK, M.; LENARČIČ, B.; TURK, V. Cloning a synthetic gene for human stefin B and its expression in *E. coli*. **FEBS Letters**, v. 239, n. 1, p. 41–44, 1988. Disponível em: [https://doi.org/10.1016/0014-5793\(88\)80541-6](https://doi.org/10.1016/0014-5793(88)80541-6)

JERALA, R.; ZEROVNIK, E.; KIDRIC, J.; TURK, V. pH-induced Conformational Transitions of the Propeptide of Human Cathepsin L. **Journal of Biological Chemistry**, v. 273, n. 19, p. 11498–11504, 1998. Disponível em: <https://doi.org/10.1074/jbc.273.19.11498>

KÄLL, L.; KROGH, A.; SONNHAMMER, E. L. . A Combined Transmembrane Topology and Signal Peptide Prediction Method. **Journal of Molecular Biology**, v. 338, n. 5, p. 1027–1036, 2004. Disponível em: <https://doi.org/10.1016/j.jmb.2004.03.016>

KEPPLER, D.; SLOANE, B. F. Cathepsin B: Multiple Enzyme Forms from a Single Gene and Their Relation to Cancer. **Enzyme and Protein**, v. 49, n. 1–3, p. 94–105, 1996. Disponível em: <https://doi.org/10.1159/000468619>

KIGGUNDU, A.; GOULET, M.-C.; GOULET, C.; DUBUC, J.-F.; RIVARD, D.; BENCHABANE, M.; PÉPIN, G.; VYVER, C. van der; KUNERT, K.; MICHAUD, D. Modulating the proteinase inhibitory profile of a plant cystatin by single mutations at positively selected amino acid sites. **The Plant Journal**, v. 48, n. 3, p. 403–413, 2006. Disponível em: <https://doi.org/10.1111/j.1365-313X.2006.02878.x>

LAEMMLI, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. **Nature**, v. 227, n. 5259, p. 680–685, 1970. Disponível em: <https://doi.org/10.1038/227680a0>

LANG, L.; REITMAN, M.; TANG, J.; ROBERTS, R. M.; KORNFELD, S. Lysosomal enzyme phosphorylation. Recognition of a protein-dependent determinant allows specific phosphorylation of oligosaccharides present on lysosomal enzymes. **Journal of Biological Chemistry**, v. 259, n. 23, p. 14663–14671, 1984. Disponível em:

[https://doi.org/10.1016/S0021-9258\(17\)42654-8](https://doi.org/10.1016/S0021-9258(17)42654-8)

LEGUIZAMÓN, N. D. P. *et al.* In vivo and in vitro anti-inflammatory and pro-osteogenic effects of citrus cystatin CsinCPI-2. **Cytokine**, v. 123, p. 154760, 2019. Disponible em: <https://doi.org/10.1016/j.cyto.2019.154760>

LI, L.; WANG, X.; FEI, X.; XIA, L.; QIN, Z.; LIANG, Z. Parkinson's disease involves autophagy and abnormal distribution of cathepsin L. **Neuroscience Letters**, v. 489, n. 1, p. 62–67, 2011. Disponible em: <https://doi.org/10.1016/j.neulet.2010.11.068>

LIANG, J.; WANG, Y.; DING, G.; LI, W.; YANG, G.; HE, N. Biotic stress-induced expression of mulberry cystatins and identification of cystatin exhibiting stability to silkworm gut proteinases. **Planta**, v. 242, n. 5, p. 1139–1151, 2015. Disponible em: <https://doi.org/10.1007/s00425-015-2345-x>

MARGIS, R.; REIS, E. M.; VILLERET, V. Structural and Phylogenetic Relationships among Plant and Animal Cystatins. **Archives of Biochemistry and Biophysics**, v. 359, n. 1, p. 24–30, 1998. Disponible em: <https://doi.org/10.1006/abbi.1998.0875>

MARTÍNEZ, M.; ABRAHAM, Z.; CARBONERO, P.; DÍAZ, I. Comparative phylogenetic analysis of cystatin gene families from arabidopsis, rice and barley. **Molecular Genetics and Genomics**, v. 273, n. 5, p. 423–432, 2005. Disponible em: <https://doi.org/10.1007/s00438-005-1147-4>

MARTINEZ, M.; CAMBRA, I.; CARRILLO, L.; DIAZ-MENDOZA, M.; DIAZ, I. Characterization of the Entire Cystatin Gene Family in Barley and Their Target Cathepsin L-Like Cysteine-Proteases, Partners in the Hordein Mobilization during Seed Germination. **Plant Physiology**, v. 151, n. 3, p. 1531–1545, 2009. Disponible em: <https://doi.org/10.1104/pp.109.146019>

MARTINEZ, M.; DIAZ-MENDOZA, M.; CARRILLO, L.; DIAZ, I. Carboxy terminal extended phytocystatins are bifunctional inhibitors of papain and legumain cysteine proteinases. **FEBS Letters**, v. 581, n. 16, p. 2914–2918, 2007. Disponible em: <https://doi.org/10.1016/j.febslet.2007.05.042>

MARTINEZ, M.; RUBIO-SOMOZA, I.; CARBONERO, P.; DIAZ, I. A cathepsin B-like cysteine protease gene from *Hordeum vulgare* (gene CatB) induced by GA in aleurone cells is under circadian control in leaves. **Journal of Experimental Botany**, v. 54, n. 384, p. 951–959, 2003. Disponible em: <https://doi.org/10.1093/jxb/erg099>

MARTINEZ, M.; SANTAMARIA, M.; DIAZ-MENDOZA, M.; ARNAIZ, A.; CARRILLO, L.; ORTEGO, F.; DIAZ, I. Phytocystatins: Defense Proteins against Phytophagous Insects and Acari. **International Journal of Molecular Sciences**, v. 17, n. 10, p. 1747, 2016. Disponível em: <https://doi.org/10.3390/ijms17101747>

MCGRATH, M. E. the Lysosomal Cysteine Proteases. **Annual Review of Biophysics and Biomolecular Structure**, v. 28, n. 1, p. 181–204, 1999. Disponível em: <https://doi.org/10.1146/annurev.biophys.28.1.181>

MCGUFFIN, L. J.; BRYSON, K.; JONES, D. T. The PSIPRED protein structure prediction server. **Bioinformatics**, v. 16, n. 4, p. 404–405, 2000. Disponível em: <https://doi.org/10.1093/bioinformatics/16.4.404>

MELO, I. R. S.; DIAS, L. P.; ARAÚJO, N. M. S.; VASCONCELOS, I. M.; MARTINS, T. F.; DE MORAIS, G. A.; GONÇALVES, J. F. C.; NAGANO, C. S.; CARNEIRO, R. F.; OLIVEIRA, J. T. A. CICPI, a cysteine protease inhibitor purified from *Cassia leiandra* seeds has antifungal activity against *Candida tropicalis* by inducing disruption of the cell surface. **International Journal of Biological Macromolecules**, v. 133, p. 1115–1124, 2019. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2019.04.174>

MELO, P. M. S.; EL CHAMY MALUF, S.; AZEVEDO, M. F.; PASCHOALIN, T.; BUDU, A.; BAGNARESI, P.; HENRIQUE-SILVA, F.; SOARES-COSTA, A.; GAZARINI, M. L.; CARMONA, A. K. Inhibition of *Plasmodium falciparum* cysteine proteases by the sugarcane cystatin CaneCPI-4. **Parasitology International**, v. 67, n. 2, p. 233–236, 2018. Disponível em: <https://doi.org/10.1016/j.parint.2017.12.005>

MELO, R. L.; ALVES, L. C.; DEL NERY, E.; JULIANO, L.; JULIANO, M. A. Synthesis and Hydrolysis by Cysteine and Serine Proteases of Short Internally Quenched Fluorogenic Peptides. **Analytical Biochemistry**, v. 293, n. 1, p. 71–77, 2001. Disponível em: <https://doi.org/10.1006/abio.2001.5115>

MIRANDA-ROTTMANN, S.; ASPILLAGA, A. A.; PÉREZ, D. D.; VASQUEZ, L.; MARTINEZ, A. L. F.; LEIGHTON, F. Juice and Phenolic Fractions of the Berry *Aristotelia chilensis* Inhibit LDL Oxidation in Vitro and Protect Human Endothelial Cells against Oxidative Stress. **Journal of Agricultural and Food Chemistry**, v. 50, n. 26, p. 7542–7547, 2002. Disponível em: <https://doi.org/10.1021/jf025797n>

MISLE, E.; GARRIDO, E.; CONTARDO, H.; GONZÁLEZ, W. Maqui [*Aristotelia chilensis*

(Mol.) Stuntz]-the Amazing Chilean Tree: A Review. **Journal of Agricultural Science and Technology**, v. B1, p. 473–482, 2011.

MONTEIRO JÚNIOR, J. E.; VALADARES, N. F.; PEREIRA, H. D.; DYSZY, F. H.; DA COSTA FILHO, A. J.; UCHÔA, A. F.; DE OLIVEIRA, A. S.; DA SILVEIRA CARVALHO, C. P.; GRANGEIRO, T. B. Expression in *Escherichia coli* of cysteine protease inhibitors from cowpea (*Vigna unguiculata*): The crystal structure of a single-domain cystatin gives insights on its thermal and pH stability. **International Journal of Biological Macromolecules**, v. 102, p. 29–41, 2017. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2017.04.008>

MORT, J. S.; BUTTLE, D. J. Cathepsin B. **The International Journal of Biochemistry & Cell Biology**, v. 29, n. 5, p. 715–720, 1997. Disponível em: [https://doi.org/10.1016/S1357-2725\(96\)00152-5](https://doi.org/10.1016/S1357-2725(96)00152-5)

MURASHIGE, T.; SKOOG, F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. **Physiologia Plantarum**, v. 15, n. 3, p. 473–497, 1962. Disponível em: <https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>

MUSIL, D.; ZUCIC, D.; TURK, D.; ENGH, R. A.; MAYR, I.; HUBER, R.; POPOVIC, T.; TURK, V.; TOWATARI, T.; KATUNUMA, N. The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. **The EMBO journal**, v. 10, n. 9, p. 2321–2330, 1991. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/1868826>

NAGASE, H.; SALVESEN, G. S. Finding, purification and characterization of natural protease inhibitors. In: BEYNON, R. J.; BOND, S. (org.). **Proteolytic Enzymes**. New York: Oxford, 2001. p. 359. *E-book*. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/030744129090038P>

NISHIMURA, Y.; KAWABATA, T.; FURUNO, K.; KATO, K. Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. **Archives of Biochemistry and Biophysics**, v. 271, n. 2, p. 400–406, 1989. Disponível em: [https://doi.org/10.1016/0003-9861\(89\)90289-0](https://doi.org/10.1016/0003-9861(89)90289-0)

NYCANDER, M.; ESTRADA, S.; MORT, J. S.; ABRAHAMSON, M.; BJÖRK, I. Two-step mechanism of inhibition of cathepsin B by cystatin C due to displacement of the proteinase occluding loop. **FEBS Letters**, v. 422, n. 1, p. 61–64, 1998. Disponível em: [https://doi.org/10.1016/S0014-5793\(97\)01604-9](https://doi.org/10.1016/S0014-5793(97)01604-9)

OHTSUBO, S.; TAIYOJI, M.; KAWASE, T.; TANIGUCHI, M.; SAITOH, E. Oryzacystatin-

II, a Cystatin from Rice (*Oryza sativa* L. japonica), Is a Dimeric Protein: Possible Involvement of the Interconversion between Dimer and Monomer in the Regulation of the Reactivity of Oryzacystatin-II. **Journal of Agricultural and Food Chemistry**, v. 55, n. 5, p. 1762–1766, 2007. Disponível em: <https://doi.org/10.1021/jf062637t>

OLIVEIRA, J.; MAGLIARELLI, H.; PEREIRA, F.; GIANOTTI, A.; SOARES-COSTA, A.; HENRIQUE-SILVA, F.; SOARES, I. C.; NONOGAKI, S.; TRAVASSOS, L. R.; CARMONA, A. K.; PASCHOALIN, T. Sugarcane Cystatin CaneCPI-4 inhibits Melanoma Growth by Angiogenesis Disruption. **Journal of Cancer Science & Therapy**, v. 03, n. 07, 2011. Disponível em: <https://doi.org/10.4172/1948-5956.1000081>

PAVLOVA, A.; KRUPA, J. C.; MORT, J. S.; ABRAHAMSON, M.; BJÖRK, I. Cystatin inhibition of cathepsin B requires dislocation of the proteinase occluding loop. Demonstration by release of loop anchoring through mutation of His110. **FEBS Letters**, v. 487, n. 2, p. 156–160, 2000. Disponível em: [https://doi.org/10.1016/S0014-5793\(00\)02337-1](https://doi.org/10.1016/S0014-5793(00)02337-1)

PELÁ, V. T. *et al.* Safety and In Situ Antierosive Effect of CaneCPI-5 on Dental Enamel. **Journal of Dental Research**, v. 100, n. 12, p. 1344–1350, 2021. Disponível em: <https://doi.org/10.1177/00220345211011590>

PERNAS, M.; SÁNCHEZ-MONGE, R.; SALCEDO, G. Biotic and abiotic stress can induce cystatin expression in chestnut. **FEBS Letters**, v. 467, n. 2–3, p. 206–210, 2000. Disponível em: [https://doi.org/10.1016/S0014-5793\(00\)01157-1](https://doi.org/10.1016/S0014-5793(00)01157-1)

QURAIISHI, O.; NÄGLER, D. K.; FOX, T.; SIVARAMAN, J.; CYGLER, M.; MORT, J. S.; STORER, A. C. The Occluding Loop in Cathepsin B Defines the pH Dependence of Inhibition by Its Propeptide. **Biochemistry**, v. 38, n. 16, p. 5017–5023, 1999. Disponível em: <https://doi.org/10.1021/bi981950o>

RAWLINGS, N. D.; BARRETT, A. J. Evolution of proteins of the cystatin superfamily. **Journal of Molecular Evolution**, v. 30, n. 1, p. 60–71, 1990. Disponível em: <https://doi.org/10.1007/BF02102453>

RAWLINGS, N. D.; BARRETT, A. J. Families of cysteine peptidases. *In*: BARRETT, A. J. (org.). **Methods in Enzymology: Proteolytic Enzymes: Serine and Cysteine Peptidases**. Cambridge: Academic Press, 1994. p. 461–486. Disponível em: [https://doi.org/10.1016/0076-6879\(94\)44034-4](https://doi.org/10.1016/0076-6879(94)44034-4)

RAWLINGS, N. D.; BARRETT, A. J.; THOMAS, P. D.; HUANG, X.; BATEMAN, A.; FINN,

R. D. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. **Nucleic Acids Research**, v. 46, n. D1, p. D624–D632, 2018. Disponível em: <https://doi.org/10.1093/nar/gkx1134>

REISER, J.; ADAIR, B.; REINHECKEL, T. Specialized roles for cysteine cathepsins in health and disease. **Journal of Clinical Investigation**, v. 120, n. 10, p. 3421–3431, 2010. Disponível em: <https://doi.org/10.1172/JCI42918>

ROBERT, X.; GOUET, P. Deciphering key features in protein structures with the new ENDscript server. **Nucleic Acids Research**, v. 42, n. W1, p. W320–W324, 2014. Disponível em: <https://doi.org/10.1093/nar/gku316>

ROJO, L. E.; RIBNICKY, D.; LOGENDRA, S.; POULEV, A.; ROJAS-SILVA, P.; KUHN, P.; DORN, R.; GRACE, M. H.; LILA, M. A.; RASKIN, I. In vitro and in vivo anti-diabetic effects of anthocyanins from Maqui Berry (*Aristotelia chilensis*). **Food Chemistry**, v. 131, n. 2, p. 387–396, 2012. Disponível em: <https://doi.org/10.1016/j.foodchem.2011.08.066>

ROSSI, A.; DEVERAUX, Q.; TURK, B.; SALI, A. Comprehensive search for cysteine cathepsins in the human genome. **Biological Chemistry**, v. 385, n. 5, 2004. Disponível em: <https://doi.org/10.1515/BC.2004.040>

SAMBROOK, J.; RUSSEL, D. **Molecular Cloning: A Laboratory Manual**. 3^a ed. Nova Iorque: Cold Spring Harbor Laboratory Press, 2000.

SANTIAGO, A. C.; KHAN, Z. N.; MIGUEL, M. C.; GIRONDA, C. C.; SOARES-COSTA, A.; PELÁ, V. T.; LEITE, A. L.; EDWARDSON, J. M.; BUZALAF, M. A. R.; HENRIQUE-SILVA, F. A New Sugarcane Cystatin Strongly Binds to Dental Enamel and Reduces Erosion. **Journal of Dental Research**, v. 96, n. 9, p. 1051–1057, 2017. Disponível em: <https://doi.org/10.1177/0022034517712981>

SCHNEIDER, V. K.; DA SILVA FERRARA, T. F.; ROCHA, S. V.; SANTOS-JÚNIOR, C. D.; NEO-JUSTINO, D. M.; DA CUNHA, A. F.; DE OLIVEIRA DA SILVA, J. P. M.; DOS SANTOS TERSARIOL, I. L.; CARMONA, A. K.; HENRIQUE-SILVA, F.; SOARES-COSTA, A. Recombinant expression, characterization and phylogenetic studies of novel cystatins-like proteins of sweet orange (*Citrus sinensis*) and clementine (*Citrus clementina*). **International Journal of Biological Macromolecules**, v. 152, p. 546–553, 2020. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2020.02.280>

SHIBAO, P. Y. T. *et al.* Sugarcane cystatins: From discovery to biotechnological applications.

International Journal of Biological Macromolecules, v. 167, p. 676–686, 2021. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2020.11.185>

SHIN, D. *et al.* Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. **Nature**, v. 587, n. 7835, p. 657–662, 2020. Disponível em: <https://doi.org/10.1038/s41586-020-2601-5>

SHU, D. J. H.; CHYAN, C.-L.; TZEN, J. T. C.; CHOU, W.-M. Molecular Cloning, Expression, and Functional Characterization of a Cystatin from Pineapple Stem. **Bioscience, Biotechnology, and Biochemistry**, v. 68, n. 8, p. 1681–1689, 2004. Disponível em: <https://doi.org/10.1271/bbb.68.1681>

SHYU, D. J. H.; YONG, Y.-M.; LU, H.-C.; CHENG, Y.-M.; TZEN, J. T. C.; CHOU, W.-M. Cloning, functional expression and characterization of a phytocystatin gene from jelly fig (*Ficus awkeotsang* Makino) achenes. **Botanical Studies**, v. 52, n. 4, p. 407–416, 2011.

SOARES-COSTA, A.; BELTRAMINI, L. M.; THIEMANN, O. H.; HENRIQUE-SILVA, F. A sugarcane cystatin: recombinant expression, purification, and antifungal activity. **Biochemical and Biophysical Research Communications**, v. 296, n. 5, p. 1194–1199, 2002. Disponível em: [https://doi.org/10.1016/S0006-291X\(02\)02046-6](https://doi.org/10.1016/S0006-291X(02)02046-6)

STORER, A. C.; MÉNARD, R. Catalytic mechanism in papain family of cysteine peptidases. *In*: BARRETT, A. J. (org.). **Proteolytic Enzymes: Serine and Cysteine Peptidases (Methods in Enzymology Series)**. San Diego: Academic Press, 1994. p. 486–500. Disponível em: [https://doi.org/10.1016/0076-6879\(94\)44035-2](https://doi.org/10.1016/0076-6879(94)44035-2)

STORER, A. C.; MÉNARD, R. Papain. *In*: RAWLINGS, N. D.; SALVESEN, G. (org.). **Handbook of Proteolytic Enzymes**. 3^a ed. London: Elsevier, 2013. p. 1858–1861. *E-book*. Disponível em: <https://doi.org/10.1016/B978-0-12-382219-2.00418-X>

SUDHAN, D. R.; SIEMANN, D. W. Cathepsin L targeting in cancer treatment. **Pharmacology & Therapeutics**, v. 155, p. 105–116, 2015. Disponível em: <https://doi.org/10.1016/j.pharmthera.2015.08.007>

SUGAWARA, H.; SHIBUYA, K.; YOSHIOKA, T.; HASHIBA, T.; SATOH, S. Is a cysteine proteinase inhibitor involved in the regulation of petal wilting in senescing carnation (*Dianthus caryophyllus* L.) flowers? **Journal of Experimental Botany**, v. 53, n. 368, p. 407–413, 2002. Disponível em: <https://doi.org/10.1093/jexbot/53.368.407>

SZYMAŃSKA, A.; RADULSKA, A.; CZAPLEWSKA, P.; GRUBB, A.; GRZONKA, Z.; RODZIEWICZ-MOTOWIDŁO, S. Governing the monomer-dimer ratio of human cystatin c by single amino acid substitution in the hinge region. **Acta biochimica Polonica**, v. 56, n. 3, p. 455–63, 2009. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19636441>

TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. **Molecular Biology and Evolution**, v. 38, n. 7, p. 3022–3027, 2021. Disponível em: <https://doi.org/10.1093/molbev/msab120>

TAN, Y.; LI, M.; YANG, Y.; SUN, X.; WANG, N.; LIANG, B.; MA, F. Overexpression of MpCYS4, A Phytocystatin Gene from *Malus prunifolia* (Willd.) Borkh., Enhances Stomatal Closure to Confer Drought Tolerance in Transgenic *Arabidopsis* and Apple. **Frontiers in Plant Science**, v. 8, 2017. Disponível em: <https://doi.org/10.3389/fpls.2017.00033>

TAYLOR, M. A. J.; BAKER, K. C.; BRIGGS, G. S.; CONNERTON, I. F.; CUMMINGS, N. J.; PRATT, K. A.; REVELL, D. F.; FREEDMAN, R. B.; GOODENOUGH, P. W. Recombinant pro-regions from papain and papaya proteinase IV are selective high affinity inhibitors of the mature papaya enzymes. “**Protein Engineering, Design and Selection**”, v. 8, n. 1, p. 59–62, 1995. Disponível em: <https://doi.org/10.1093/protein/8.1.59>

TRENTINI, G. E.; ROJAS, M.; GAJARDO, D.; ALBURQUENQUE, D.; VILLAGRA, E.; GÓMEZ, A.; ARRU, L.; ARENCIBIA, A. D. Elicitation of phenylpropanoids in maqui (*Aristotelia chilensis* [Mol.] Stuntz) plants micropropagated in photomixotrophic temporary immersion bioreactors (TIBs). **Plant Cell, Tissue and Organ Culture (PCTOC)**, v. 146, n. 3, p. 607–619, 2021. Disponível em: <https://doi.org/10.1007/s11240-021-02097-0>

TURK, B.; TURK, D.; TURK, V. Lysosomal cysteine proteases: more than scavengers. **Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology**, v. 1477, n. 1–2, p. 98–111, 2000. Disponível em: [https://doi.org/10.1016/S0167-4838\(99\)00263-0](https://doi.org/10.1016/S0167-4838(99)00263-0)

TURK, V. NEW EMBO MEMBERS’ REVIEW: Lysosomal cysteine proteases: facts and opportunities. **The EMBO Journal**, v. 20, n. 17, p. 4629–4633, 2001. Disponível em: <https://doi.org/10.1093/emboj/20.17.4629>

TURK, V.; BODE, W. The cystatins: Protein inhibitors of cysteine proteinases. **FEBS Letters**, v. 285, n. 2, p. 213–219, 1991. Disponível em: [https://doi.org/10.1016/0014-5793\(91\)80804-C](https://doi.org/10.1016/0014-5793(91)80804-C)

TURK, V.; STOKA, V.; TURK, D. Cystatins: Biochemical and structural properties, and

medical relevance. **Frontiers in Bioscience**, v. Volume, n. 13, p. 5406, 2008. Disponível em: <https://doi.org/10.2741/3089>

TURK, V.; STOKA, V.; VASILJEVA, O.; RENKO, M.; SUN, T.; TURK, B.; TURK, D. Cysteine cathepsins: From structure, function and regulation to new frontiers. **Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics**, v. 1824, n. 1, p. 68–88, 2012. Disponível em: <https://doi.org/10.1016/j.bbapap.2011.10.002>

VALADARES, N. F.; DELLAMANO, M.; SOARES-COSTA, A.; HENRIQUE-SILVA, F.; GARRATT, R. C. Molecular determinants of improved cathepsin B inhibition by new cystatins obtained by DNA shuffling. **BMC Structural Biology**, v. 10, n. 1, p. 30, 2010. Disponível em: <https://doi.org/10.1186/1472-6807-10-30>

VALADARES, N. F.; OLIVEIRA-SILVA, R.; CAVINI, I. A.; ALMEIDA MARQUES, I.; D’MUNIZ PEREIRA, H.; SOARES-COSTA, A.; HENRIQUE-SILVA, F.; KALBITZER, H. R.; MUNTE, C. E.; GARRATT, R. C. X-ray crystallography and NMR studies of domain-swapped canecystatin-1. **The FEBS Journal**, v. 280, n. 4, p. 1028–1038, 2013. Disponível em: <https://doi.org/10.1111/febs.12095>

VALDES-RODRIGUEZ, S.; CEDRO-TANDA, A.; AGUILAR-HERNANDEZ, V.; CORTES-ONOFRE, E.; BLANCO-LABRA, A.; GUERRERO-RANGEL, A. Recombinant amaranth cystatin (AhCPI) inhibits the growth of phytopathogenic fungi. **Plant Physiology and Biochemistry**, v. 48, n. 6, p. 469–475, 2010. Disponível em: <https://doi.org/10.1016/j.plaphy.2010.03.012>

VALDÉS-RODRÍGUEZ, S.; GUERRERO-RANGEL, A.; MELGOZA-VILLAGÓMEZ, C.; CHAGOLLA-LÓPEZ, A.; DELGADO-VARGAS, F.; MARTÍNEZ-GALLARDO, N.; SÁNCHEZ-HERNÁNDEZ, C.; DÉLANO-FRIER, J. Cloning of a cDNA encoding a cystatin from grain amaranth (*Amaranthus hypochondriacus*) showing a tissue-specific expression that is modified by germination and abiotic stress. **Plant Physiology and Biochemistry**, v. 45, n. 10–11, p. 790–798, 2007. Disponível em: <https://doi.org/10.1016/j.plaphy.2007.07.007>

VAN DER VYVER, C.; SCHNEIDEREIT, J.; DRISCOLL, S.; TURNER, J.; KUNERT, K.; FOYER, C. H. Oryzacystatin I expression in transformed tobacco produces a conditional growth phenotype and enhances chilling tolerance. **Plant Biotechnology Journal**, v. 1, n. 2, p. 101–112, 2003. Disponível em: <https://doi.org/10.1046/j.1467-7652.2003.00010.x>

VERMA, M.; CHOI, J.; COTTRELL, K. A.; LAVAGNINO, Z.; THOMAS, E. N.;

PAVLOVIC-DJURANOVIC, S.; SZCZESNY, P.; PISTON, D. W.; ZAHER, H. S.; PUGLISI, J. D.; DJURANOVIC, S. A short translational ramp determines the efficiency of protein synthesis. **Nature Communications**, v. 10, n. 1, p. 5774, 2019. Disponível em: <https://doi.org/10.1038/s41467-019-13810-1>

VIDAK, E.; JAVORŠEK, U.; VIZOVIŠEK, M.; TURK, B. Cysteine Cathepsins and their Extracellular Roles: Shaping the Microenvironment. **Cells**, v. 8, n. 3, p. 264, 2019. Disponível em: <https://doi.org/10.3390/cells8030264>

VON HEIJNE, G. Signal sequences. **Journal of Molecular Biology**, v. 184, n. 1, p. 99–105, 1985. Disponível em: [https://doi.org/10.1016/0022-2836\(85\)90046-4](https://doi.org/10.1016/0022-2836(85)90046-4)

WANG, X.; CHAN, S. J.; EDDY, R. L.; BYERS, M. G.; FUKUSHIMA, Y.; HENRY, W. M.; SHOWS, T. B. Chromosome Assignment Of Cathepsin-B (Ctsb) To 8p22 And Cathepsin-H (Ctsh) To 15q24-q25. **Cytogenetics and Cell Genetics**, v. 46, n. 1–4, p. 710–711, 1987.

WHELAN, S.; GOLDMAN, N. A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. **Molecular Biology and Evolution**, v. 18, n. 5, p. 691–699, 2001. Disponível em: <https://doi.org/10.1093/oxfordjournals.molbev.a003851>

WIEDERANDERS, B. Structure-function relationships in class CA1 cysteine peptidase propeptides. **Acta Biochimica Polonica**, v. 50, n. 3, p. 691–713, 2003. Disponível em: https://doi.org/10.18388/abp.2003_3661

WILLSTÄTTER, R.; BAMANN, E. Über die Proteasen der Magenschleimhaut. Erste Abhandlung über die Enzyme der Leukocyten. **Hoppe-Seyler's Zeitschrift für physiologische Chemie**, v. 180, n. 1–3, p. 127–143, 1929. Disponível em: <https://doi.org/10.1515/bchm2.1929.180.1-3.127>

ZHANG, X.; LIU, S.; TAKANO, T. Two cysteine proteinase inhibitors from Arabidopsis thaliana, AtCYSa and AtCYSb, increasing the salt, drought, oxidation and cold tolerance. **Plant Molecular Biology**, v. 68, n. 1–2, p. 131–143, 2008. Disponível em: <https://doi.org/10.1007/s11103-008-9357-x>

APÊNDICE A – Produção da *Papain like protease* de SARSCoV-2

Um dos objetivos do trabalho era testar o potencial das cistatinas de maqui frente a Papain-like protease (PLPro) de SARSCoV-2. Para isso, foi proposto realizar a clonagem da sequência codificadora dessa proteína para expressão recombinante, e em seguida a realização de ensaios fluorimétricos.

Estudos prévios haviam reportado o potencial inibitório das cistatinas C e D humanas frente a alguns vírus da família Coronaviridae, quando em cultura celular. Havia uma crença pelos pesquisadores de que o mecanismo estaria envolvido com uma proteína semelhante a papaína (PLPro), que estaria sujeita a inibição por esses inibidores (COLLINS; GRUBB, 1991, 1998). No entanto, ao realizarmos ensaios fluorimétricos de inibição da PLPro de SARSCoV-2 com cistatinas de plantas, observou-se que não havia indício de inibição. Um dos fatores que poderiam ser responsáveis pela ausência de inibição, é o fato de que embora seja denominada da “semelhante a papaína”, a PLPro não pertence à família C1A das cisteíno peptidases, mas à C16 (RAWLINGS *et al.*, 2018). No entanto, a inibição da sobrevivência do vírus em cultura poderia estar associada à outra cisteíno peptidase, a catepsina L do hospedeiro, que está associada a entrada do vírus na célula (GOMES *et al.*, 2020).

Embora os ensaios de atividade não tenham apresentado sucesso, a partir dessa proposta foi possível colaborar no desenvolvimento do trabalho intitulado “*Non-Toxic Dimeric Peptides Derived from the Bothropstoxin-I Are Potent SARS-CoV-2 and Papain-like Protease Inhibitors*” (FREIRE *et al.*, 2021), com a construção do plasmídeo.

A expressão e purificação da *Papain-like protease*, foi realizada conforme previamente descrito (SHIN *et al.*, 2020), mas de maneira adaptada. A sequência de nucleotídeos de 4954 a 5902 do genoma de SARS-CoV-2 (GenBank: MT126808.1), foi a região selecionada para clonagem pois codifica os aminoácidos 1564 a 1880 (GenBank: QIG55993.1) da poliproteína que contém a PLPro. O RNA viral inativado de SARS-CoV-2 foi gentilmente cedido pelo Professor Dr. Glaucius Oliva, do Instituto de Física da USP, São Carlos, a partir do qual a sequência foi retrotranscrita e amplificada. Para isso, foram utilizados os *primers* F: 5'-ATTCCATGGGCGAAGTGAGGACTATTAAGGTGTTTAC-3' e R: 5'-ATTGCTCGAGTGGTTTTATGGTTGTTGTGTA ACT-3' que contém sítios de restrição para *NcoI* e *XhoI*, em negrito. Os *amplicons* purificados assim como o plasmídeo pET-28a (Novagen) foram clivados com as enzimas *NcoI* e *XhoI* e ligados entre si, conforme descrito na seção 3.2.

Células de *E. coli* Rosetta (DE3) transformadas com o pET28a-PLpro, foram inoculadas em 5 mL de meio LB contendo cloranfenicol e canamicina a 25 µg/mL *overnight*, a 37°C e 250 rpm. O pré-inóculo foi diluído em 500 mL do mesmo meio de cultura nas mesmas condições de temperatura e agitação até atingir DO₆₀₀ de 1, quando foi adicionado ao meio IPTG (5 mM) e ZnCl₂ (1 mM). A indução foi conduzida *overnight* a 18°C. Após esse período, foi realizada a separação da fração solúvel conforme a seção 3.3, porém com um tampão de lise composto por Tris-HCl 50 mM, NaCl 150 mM, imidazol 10 mM, DTT 2 mM, pH 8,5.

A purificação por coluna de afinidade foi executada conforme descrito na seção 3.3. As frações selecionadas foram dialisadas com o tampão de armazenamento (Tris-HCl 20 mM, NaCl 100 mM, DTT 2 mM, pH 8,5) em três ciclos de 1 hora.

A produção heteróloga da proteína foi efetuada com sucesso, apresentando uma quantidade considerável na fração solúvel, o que permitiu a purificação pela coluna (Figura 31).

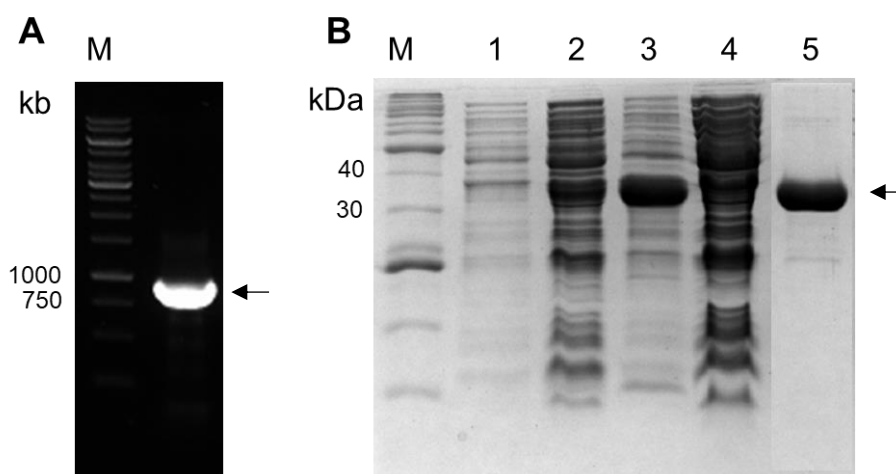


Figura 31: Clonagem e expressão da PLPro de SARSCoV-2. (A) Produto de amplificação (indicado pela seta). M: Marcador de peso molecular (GeneRuler 1 kb ThermoScientific) (B) Análise em SDS-PAGE 15 % das amostras coletadas antes (1) e ao término da indução (2), além das frações insolúvel (3) e solúvel (4) e purificada (5). M: BenchMark Protein Ladder – ThermoFisher Scientific. A banda referente a PLPro está indicada pela seta.

Os ensaios de atividade foram conduzidos de maneira semelhante à seção 3.4, com algumas adaptações. Resumidamente, a PLPro (1,27 µM) foi pré-incubada em tampão (Tris 20 mM, NaCl 5 mM, DTT 5 mM, pH 8,0) durante cinco minutos. Após isso, foi adicionado o substrato Z-LRGG-AMC (0,6 µM) e subsequentemente concentrações crescentes de cistatinas. No entanto, não foi possível observar qualquer resultado, que pudesse afirmar que houvesse inibição.

Artigo publicado referente ao tema:

Marjorie C. L. C. Freire, Gabriela D. Noske, Natália V. Bitencourt, Paulo R. S. Sanches, Norival A. Santos-Filho, Victor O. Gawriljuk, **Eduardo P. de Souza**, Victor H. R. Nogueira,

Mariana O. de Godoy, Aline M. Nakamura, Rafaela S. Fernandes, Andre S. Godoy, Maria A. Juliano, Bianca M. Peres, Cecília G. Barbosa, Carolina B. Moraes, Lucio H. G. de Freitas-Junior, Eduardo M. Cilli, Rafael V. C. Guido, Glaucius Oliva. Non-toxic dimeric peptides derived from the Bothropstoxin-I are potent SARS-CoV-2 and Papain-Like Protease inhibitors. *Molecules*. v. 26, n. 16, p. 4896, 12 ago. 2021. <https://doi.org/10.3390/molecules26164896>

APÊNDICE B – Sequências das MaquiCPIs

- MaquiCPI-1

- cDNA

ATGCACCACCTTAATTTGCGTGTCCAAACACCACCCCCATTCTATAAATAAATAATAACC
 ATTCTCCTCATTAATATTCACAAAACTCTAAATCTAAATCTAAATCAAAAGCCATGGTAAAA
 CCTCTAGTAATATTGTCAGTTTGTGGATTGTATCAATTTGTTTCTGTGGTGTGGAGGGTTTT
 GGGTCAATAGTGGGAGGGAGAACGGAGGTGAAAGATGTAAAGACCAATGAGGAGGTGCA
 AGAACTGGGGAGGTTTTCGGTAGAGGAATTCAATAGAAACCAGATTCACCGCATGAGCAG
 CAATGGTGGCCAATTGTTGATGTTCTCTCAGGTTATAGAGGCGCAAAAGCAGGTGGTTTCT
 GGGATCAAGTATTTTCTCAAGATTCAAGCCATGAAAAATGGACAAACGCGTTTGTGGATT
 CGGTGGTTGTGGTTAAGCCTTGGATTCAATAATCCAAGGAGTTGCTTCATTTTCGCACCTTCT
 ACCGATTAATATTAGATTAGATTAAGGTTAATTTTATGGTATGTCAATGTTAATCTGTGGT
 TAGGTATGGTGGTGTGTATTAATCTCCTGTTGTTTTGCTTCTTTGTTAATGTTTCCTACTTT
 TGTTTTTTGTGTTGGGAATGTGTTTGTAAACCATGTTTATGTGTCCTCGAACATGTAAGGGAA
 TAAATAAACATTAATTTTAAAAAAAAAAAA

Amarelo: UTR 5'

Azul: Sequência codificadora do peptídeo sinal

Cinza: UTR 3'

- Estrutura primária da proteína

MVKPLVILSVCLIVSICFCGVEGFGSIVGGRTTEVKDVKTNEEVQELGRFSVEEFNRNQIHRMSSN
 GGQLLMFSQVIEAQKQVVSIGIKYFLKIQAMKNGQTRLFDSVVVVKPWIHNSKELLHFAPSTD

Azul: Sequência codificadora do peptídeo sinal

- Proteína de fusão com a Histag

MM: 14,45 kDa

MGSSHHHHHSSGLVPRGSHMGVEGFGSIVGGRTTEVKDVKTNEEVQELGRFSVEEFNRNQIHR
 MSSNGQLLMFSQVIEAQKQVVSIGIKYFLKIQAMKNGQTRLFDSVVVVKPWIHNSKELLHFAP
 STD

Sublinhado: Sequência de histidina

Cinza: fragmento oriundo do pET28-a

- MaquiCPI-2

-cDNA

AGATCAAAAACACCCATATGGCTACTAGCGTGCAACCAGTTTTGCCCAATTGTTGGATTAT
 AAGGCA GAGCAACCTCTTCTGTTTAATGGAGCTATAAGTGCATGGAAAGATCAAATGCCC
 GGTGGGTGGACCCTAATAAAGGACGTGAGCAAGCCACACGTGAAAGAGATCGGTGAGTTT
 GCAGTTGATGAGTATAATAAGAGATCAAAGGCTGCGTTGAAGTTTAAGACCGTAATTTGTG
 GATGGACTCAGGTGGTTTCTGGGATCAATTATAAGCTTATTGTTGAAGCTAAAGATGGAGC
 GGATACAAATTATTATGAAGCTGTAGTGTGGGAGAAAGCTTGGCAGAAGTTTAAGGAACT
 TACTTCTTTAAGGTTGTCATTGGATAA TGACATAATACTGGGCCAGCAGTGAGCATGCC
 AGATGTCACCTTAGAAGCGTATACTAAATAAATGGCTTGGTGAATGCCTCTTAGTTAAAA
 GTATGTTTTATGTATGCATGCATACGTTGGAGATGAAGATAATAAGCTCGTTGGTGTCAAG
 GTTTCAGCCTGTGGTTCGGCATTATGTATGTATACATTGATAAATGTGTACTTGTAAAGTA
 CTGTATTATTATGTTAATGAAACCTTCTTTTCGGATGGATAGCACTGGAATTAATATTTAT
 GTTGTGAGTGGAAGGTGATAA

Amarelo: UTR 5'

Azul: Sequência codificadora do peptídeo sinal

Cinza: UTR 3'

- Estrutura primária da proteína

MATSVQVPLPQLLDYKAEQPLLFNGAISA WKDQMPGGWTLIKDVSKPHVKEIGFAVDEY
 NKRSKAALKFKTVICGWTQVVS GINYKLIVEAKDGADTNYYEAVVWEKAWQKFELTSFK
 VVIG

Azul: Sequência codificadora do peptídeo sinal

- Proteína de fusão com a Histag

MM: 13,18 kDa

MGSSHHHHHSSGLVPRGSHM WKDQMPGGWTLIKDVSKPHVKEIGFAVDEYNKRSKAALK
 FKTVICGWTQVVS GINYKLIVEAKDGADTNYYEAVVWEKAWQKFELTSFKVVIG

Sublinhado: Sequência de histidina

Cinza: fragmento oriundo do pET28-a

- MaquiCPI-3

-cDNA

CTATATAAATCCACTCACCCCACCACAACCAAAAACCGCAACCTAAAAAAACACTCCAAAA
 AACAAACAGTTACAATGAAATCACTTACCCTTCTTCTAGTCATTTTCGTCACTATTTCTCTCTC
 TCTTCGCTGCCAGTGCAAGGATAGATCCTCTTGTTCGGTGGGTGGACCCCAATAACGAACAT
 GGAAGACCCGCACGTGAAAGAGATTGGTGAGTTTGCAGTTGATGCGTACAACAAGGGTTC
 AAAGGCTGCGTTGAAGTTTGGAGACCGTAATTAGCGGAGAGACTCAAGTGGTTTCTGGCAC
 CAACTATAAGCTTATTGTGGCAGCTAAAGATGGAGAAGCGGTTTCAAAGAATTATGAGGC
 TGTCGTGTGGGAGAAACCATGGGTGAATTTCAAGAATCTTACTTCTTTTAAAGTTGTGTAA
 TATGGTGATCTTTTGTAGTAATCTAAGTATTATATAAAAAGAATTGAGAATAATAAGAGT

Amarelo: UTR 5'

Azul: Sequência codificadora do peptídeo sinal

Cinza: UTR 3'

- Estrutura primária da proteína

MKSLTLLLVISLFLSLFAASARIDPLVGGWTPITNMEDPHVKEIGFAVDAYNKGSKAALKFE
 TVISGETQVVSGETNYKLIVAAKDGEAVSKNYEAVVWEKPWVNFKNLTSFKVV

Azul: Sequência codificadora do peptídeo sinal

- Proteína de fusão com a Histag

MM: 12,6 kDa

MGSSHHHHHSSGLVPRGSHMRIDPLVGGWTPITNMEDPHVKEIGFAVDAYNKGSKAALKF
 ETVISGETQVVSGETNYKLIVAAKDGEAVSKNYEAVVWEKPWVNFKNLTSFKVV

Sublinhado: Sequência de histidina

Cinza: fragmento oriundo do pET28-a

- MaquiCPI-4

-cDNA

GACTAAAAGACTTTAACACGTAAAAAATTGCTTTTTGTCTTTTACTCTCTCTCTAATATCAA
 ATATGGTTGTATATAAAAAACAGAGGCACAGGCTCGCAATTGGCATTGAGGAAGAGAAAAG
 GACCAAAGAAATAAAAGAGAAAAAAATGGCAAGAGTGGGTGGTGTAGCGATGTTGAGG
 GAAGTGCCAACAGTCTGGAGATTGATAGCCTCGCTCGTTTCGCTGTTGATGAGCACAAACA
 GAAACAGAATGCGCTTCTGGAGTTTAAGAAAGCGGTGAATGTGAAAAAGCAAGTGGTTTC
 TGGTACCATATACTACATAACCTTGGAGGCAATCGATGGTGGTCAGAAGAAAGTTTATGA
 AGCCAAGGTCTGGGATAAGCCCTGGATGAACCTCAAGGAGTTGCAGGAGTTCAAGCTTGT
 TGGTGATGCCCCCTCCGACTCCACTGCTTAGCTAGGTGACTGAAGATCGCATGTATATCC
 CTTGGTTCCAGTTGAAAAAAGAAGCGTATGAGAAGGGTGAGACTTGTCTGTTTGAAATAA
 ATGTAACAAAATTATTTCTGAATTTGGATTTTAACCTTACAATATTGTTTAGAATGTGTGGA
 TTCATCTTTTCCTTTTCATGGTACAATTCTCAAGCTTCTTTAGAACAGTGCGTACCGTGAAG
 TATGTTTTTGATATATGTTTGTCTGTTTTCACATATTTAACTGTACTACCTGAAATGCACAT
 GCTTTGTATGTGAAGATCTGTTCAACAGAAAAACTAA

Amarelo: UTR 5'

Cinza: UTR 3'

- Estrutura primária da proteína

MARVGGVSDVEGSANSLEIDSLARFAVDEHNKKQNALLEFKKAVNVKKQVVSgtIYYITL
 AIDGGQKKVYEAKVWDKPWMNFKELQEFKLVGDAPSDSTA

- Proteína de fusão com a Histag

MM: 13,36 kDa

MGSSHHHHHSSGLVPRGSHM MARVGGVSDVEGSANSLEIDSLARFAVDEHNKKQNALLEFK
 KAVNVKKQVVSgtIYYITL AIDGGQKKVYEAKVWDKPWMNFKELQEFKLVGDAPSDSTA

Sublinhado: Sequência de histidina

Cinza: fragmento oriundo do pET28-a

- MaquiCPI-5

-cDNA

AATAAATAAATAAAGAGTGTCTGAGATACGAATTCTCTCTCTAATTATCCTACTACACTCTT
 CTGCGCTTTTGACAATTTCACTTGTCTTCTCTGCTTTTGGATATTTACGCATTTGCAAACAA
 GAAATGGCCACCACCACCACCGTTCGGAGGATTGCACGAATCCCAAGGAATCAGAACAGT
 GTCGAGATCGACGGCCTCGCTCGTTTCGCGGTTGAAGAACATAACAAGAAAGAGAATAGA
 ATCCTGGAGTTTGCAAGGGTAGTAAAGGCGCAAGAACAGGTAGTTGCTGGTACTTTGCATC
 ATCTTACGGTCGAGGTTATTGACGCGGGTAAGAAGAAGCTCTATGAAGCAAAGGTTTGGG
 TGAAGCCGTGGTTGAACTTTAAAGAATTGCAAGAATTCAAGCACGTTGGTGATGTCCCATC
 TTTTACTTCTTCAGATCTTGGTGTTAAGAAAGATGGCCATGGCCCTGGATGGCAATCTGTG
 CCAGCACATGATCCTGAAGTCCAGGGTGCAGCAAATCATGCTATCAAGACCATCCAGCAG
 AGGTCCAATCATTGTATCCTTACGAACTTCAAGAAATCGTCCATGCAAACGCTGAGGTGA
 TAGAAGATTCTGCAAAATTTGATATGCTTCTCAAGGTCAAGAGGGGAGATAAAGAAGAGA
 AGCTCAAGGTTGAGGTGCACAAGAACAATGAAGGTAATTTCCATCTGAATCAGATGGAGC
 CGGACCACTCCTGATCTCACAGATTGTTGTGATGTGTCTAGATTTCTATGTGCCAAGTATCC
 TATAGTCTTTTGTGCGCAACGAACCTTTTATCTGTGCTGTAATGCTTCGCTTATACAATGTGTT
 GTATAAGACTTGTGAATCACTATTTACATAATGCTAGTCAAACACTACAGTCATGGCACATCA
 GCGTTTCAGGTGAAATGCTTGAATATGGTTATTTCAAAAATCACAAATTGTCTTGGCTTGGT
 AGTAATATTATTCTGTGAGATCGAAAGAGCGTC

Amarelo: UTR 5'

Cinza: UTR 3'

- Estrutura primária da proteína

MATTTTVGGLHESQGTQNSVEIDGLARFAVEEHNKKENRILEFARVVKAQEQVVAGTLHHL
 TVEVIDAGKKKLYEAKVWVKPWLNFKELQEFKHVGDVPSFTSSDLGVKKDGHGPGWQSV
 AHDPEVQGAANHAIKTIQQRNSLYPYELQEIVHANAIEVIEDSAKFDMLLKVKRGDKKEE
 LKVEVHKNNEGTFHLNQMEPDHS

- Proteína de fusão com a Histag

MM: 25,06 kDa

MGSSHHHHHHSSGLVPRGSHM MATTTTVGGLHESQGTQNSVEIDGLARFAVEEHNKKENRIL
 EFARVVKAQEQVVAGTLHHLTVEVIDAGKKKLYEAKVWVKPWLNFKELQEFKHVGDVPSFT
 SSDLGVKKDGHGPGWQSVPAHDPEVQGAANHAIKTIQQRNSLYPYELQEIVHANAIEVIEDSA
 KFDMLLKVKRGDKKEEKLKVEVHKNNEGTFHLNQMEPDHS

Sublinhado: Sequência de histidina

Cinza: fragmento oriundo do pET28-a

- MaquiCPI-6

-cDNA

**FCTCCAAGCAAAATCAGAGGAAGATATCTCTATTTTCATTCTAATCTTTTTTGTCTCTCAAG
TTTGTTGACACGAGGCGGAAGCTGGTTTTCGCCACGTCAAACGAATTGATTGATCTGTTC
ACCCTTTGCCTTACCAACTAAGTAACCAACTAACATTCATCGACCAACCGACACGCTTCAA
GTAACCAAAAGAATCCAAATATAGTCCAGTCCGTCTTTGGTTGTGAAAAAGAACAAAAGT
GTCAAAGTGTCAAAGTGTTTTACAGAGGAAGATGAATAAGTTGTCTGTGATAATCTTAA
GTGTTGTAGTTTTGGTTTTGTGGTTTTGTTGAATTGGTCGTTGCAACGAAGACTTGATAAA
GATGAACTTGGAGGCGTTCACGATTGCAAAGGCACTCAAACAGCGCCGAGATCGAAAG
CCTCGCTCGTTTCGCTGTCCAAGAACAACAACAAGAGAATTCACCTTCTTGAGTTTGA
AGGGTGTGAAAGGCCAAAGAACAAGTTGTTGCTGGTACGATGTACCATCTTACACTTGAAG
CAATTGATGCTGGCAAGAAGAAGATATATGAAGCTAAAGTCTGGGTAAAGCCATGGATGA
ACTTCAAGCAGTTGCAGGAATTCAAACATGCTCAGGATGGCCCTCCGTCACCACTTCAA
CCTCGGCGTTAAACCAGATGGCCATGGACTAGGATGGCAAGCAGTGTCAACAAATGATCC
TGAGGTAGAAGATGCAGCAAACCATGCTGTTAAATCCATCCAGCAGAGGTCCAACCTCATT
GTCACCCTATGAACCTTAGAGATCCTTCTAGCTAAGGCCAAGGTCATTGAAGATTATGCC
AAATTTGAATTGCTTCTGAAGGTGAGGAGGGGAATTAAGAGGAGCAGTTCAGAGTTGAA
GTAACATAAGAATAATGAAGGAAAGTTTTATATGAATTAGAATGAACTACTGCCACTGGA
TTCAGTAATTAAGTTACTATATTAAGCATCTATGGATTGAGGCTGTGATCTACTACAAT
GTGATGTAAATAATGAATAAACCCACTTAAACGATATGCATTATGTGTAAATATGCAAAT
TGCAATGTTTGCTGGCTGACTTTTGTGCGATCTAAATACAGCTACATCCTCCTTATATATTA
GATGTGTATATTATATAACAGGCTATTGTACTCTCAATTCATATTTTTTAATTCATCATGT
TGTCAGATGATGCTACT**

Amarelo: UTR 5'

Azul: Sequência codificadora do peptídeo sinal

Cinza: UTR 3'

- Estrutura primária da proteína

MNKLSVILSVVVLVCGFVELGRCNEDLIKMKLGGVHDCKGTQNSAEIESLARFAVQEHNKE
NSLLEFARVLKAKEQVVAGTMYHLTLEAIDAGKKKIYEAKVWVKPWMNFKQLQEFKHAQDG
PSVTTSNLGVKPDGHGLGWQAVSTNDPEVEDAANHAVKSIQQRNSLSPELLEILLAKAKVIE
DYAKFELLLKVRRIKKEEQFRVEVTKNNEGKFYMN

Azul: Sequência codificadora do peptídeo sinal

- Proteína de fusão com a Histag

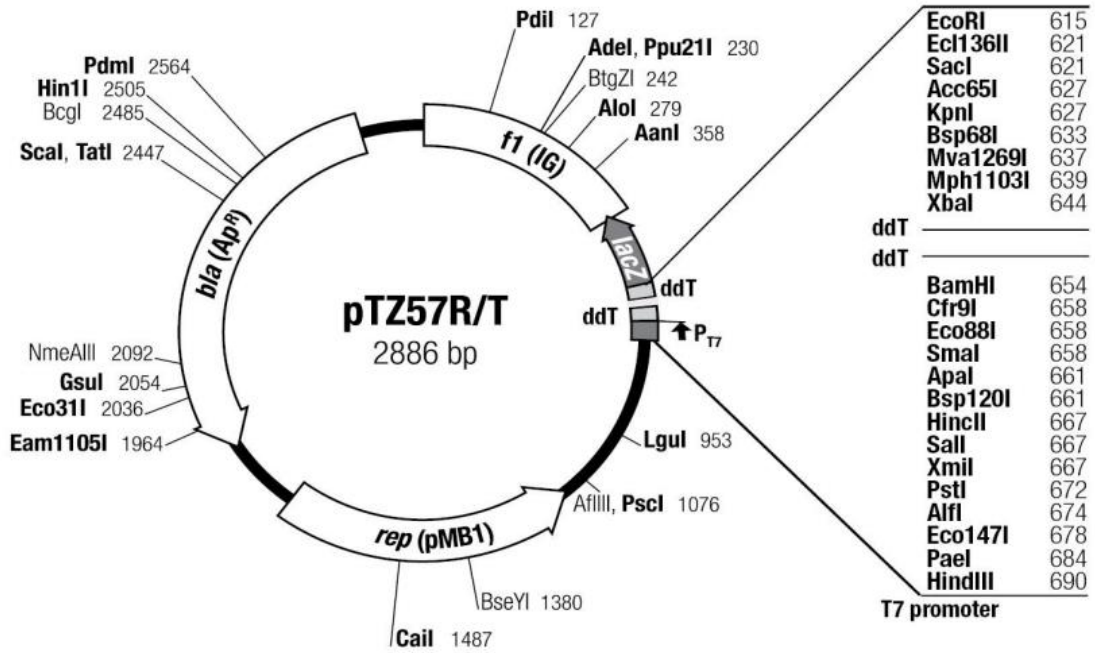
MM: 25,11 kDa

MGSSHHHHHSSGLVPRGSHMGRCNEDLIKMKLGGVHDCKGTQNSAEIESLARFAVQEHNK
ENLLEFARVLKAKEQVVAGTMYHLTLEAIDAGKKKIYEAKVWVKPWMNFKQLQEFKHAQD
GPSVTTSNLGVKPDGHGLGWQAVSTNDPEVEDAANHAVKSIQQRNSLSPELLEILLAKAKV
IEDYAKFELLLKVRRIKKEEQFRVEVTKNNEGKFYMN

Sublinhado: Sequência de histidina

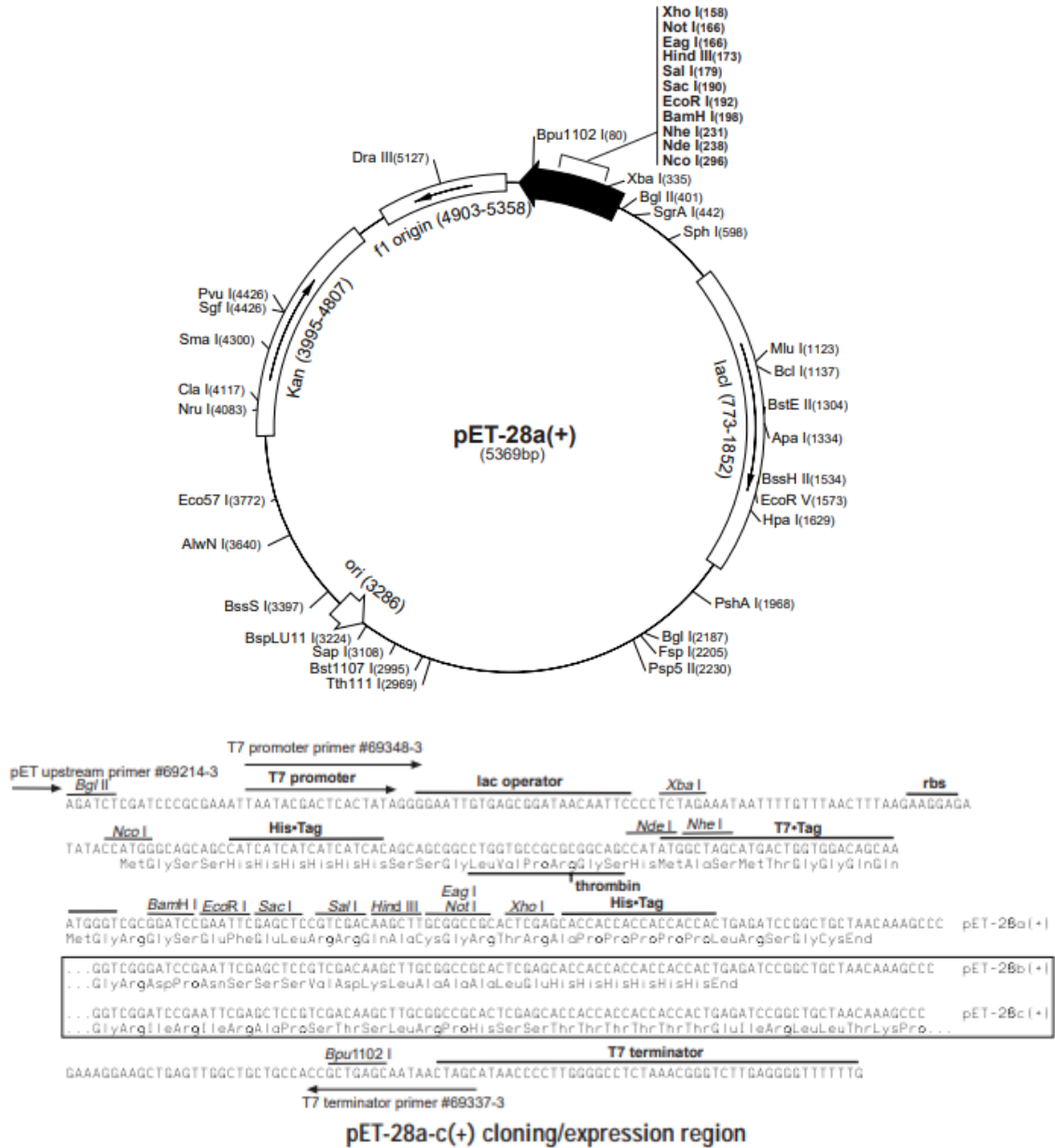
Cinza: fragmento oriundo do pET28-a

ANEXO 1 – Mapa do vetor pTZ57R_T



Fonte: Thermo Scientific.

ANEXO 2 – Mapa do vetor pET28a



Fonte: Fermentas.