

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**  
**Centro de Ciências Biológicas e da Saúde**  
**Departamento de Genética e Evolução**  
**Laboratório de Biologia Molecular**

**ANA LÍGIA BUZOLIN**

**PRODUÇÃO RECOMBINANTE E CARACTERIZAÇÃO DE UMA  
LEGUMAÍNA DE CANA-DE-AÇÚCAR**

**SÃO CARLOS**

**2014**

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

**Centro de Ciências Biológicas e da Saúde**

**Departamento de Genética e Evolução**

**Laboratório de Biologia Molecular**

**ANA LÍGIA BUZOLIN**

**PRODUÇÃO RECOMBINANTE E CARACTERIZAÇÃO DE UMA  
LEGUMAÍNA DE CANA-DE-AÇÚCAR**

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 Silva**

**SÃO CARLOS**

**2014**

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária da UFSCar**

B992pr Buzolin, Ana Lígia.  
Produção recombinante e caracterização de uma  
legumaina de cana-de-açúcar / Ana Lígia Buzolin. -- São  
Carlos : UFSCar, 2014.  
52 f.

Dissertação (Mestrado) -- Universidade Federal de São  
Carlos, 2014.

1. Biologia molecular. 2. Cisteíno peptidase. 3. Cistatina.  
4. Legumaina. 5. *Pichia pastoris*. 6. *Saccharum sp.* I. Título.

CDD: 574.88 (20ª)

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA  
MOLECULAR


“Produção recombinante e caracterização de uma  
legumaina de cana-de-açúcar”

Dissertação de Mestrado de

**ANA LÍGIA BUZOLIN**

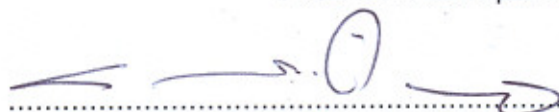
**Banca Examinadora**

Prof. Dr. Flavio Henrique Silva



.....

Prof. Dr. Henrique Ferreira



.....

Profa. Dra. Ana Paula Ulian de Araújo



.....

SÃO CARLOS – SP  
2014

*Dedico este trabalho aos meus pais,  
Paulo e Kátia, pelo amor incondicional,  
incentivo e apoio.*

*Essa conquista não seria possível sem  
vocês.*

*“Agir, eis a inteligência verdadeira. Serei o que quiser. Mas tenho que querer o que for. O êxito está em ter êxito, e não em ter condições de êxito. Condições de palácio tem qualquer terra larga, mas onde estará o palácio se não o fizerem ali?”*

*(Fernando Pessoa)*

## AGRADECIMENTOS

À minha família, especialmente meus pais e minha irmã, pelo apoio, paciência, confiança e incentivo.

Ao Prof. Dr. Flavio Henrique Silva pela orientação e pelos ensinamentos ao longo desses anos.

À Profa. Dra. Andrea Soares Costa Fuentes, por todo o aprendizado que me proporcionou e por sua disposição em ajudar sempre que precisei.

Ao MSc. Wesley Cardoso Generoso, pela valiosa amizade, pelos ensinamentos e pela ajuda desde o início, quando me mudei para São Carlos.

Às minhas amigas-irmãs Mariana e Dany, pela maravilhosa e animada companhia em todos esses anos, nas horas boas e ruins e também nos momentos de gordice. Pelo incentivo e ajuda dentro e fora do laboratório e por tornar meu dia-a-dia mais alegre.

À minha querida amiga Adelita Carolina Santiago por todo o suporte e auxílio no laboratório e pelos bons momentos fora dele, pela ótima companhia e amizade.

Aos colegas e amigos de laboratório pela convivência, ajuda e experiência compartilhada, além das descontraídas “horas do café”: Augusto, Beth, Bruna, Bruno, Célio, Daniela, Danilo, Darlan, Fernando, Gabriel, Leo, Kesser, Kishi, Marina, Naiara, Rafael, Rodrigo, Priscila, Taíse e Vanessa.

Aos amigos de Chapadão do Sul, que apesar da distância sempre me incentivaram.

Aos amigos são-carlenses, que tive o prazer de conhecer e se tornaram especiais para mim.

Ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular pelo auxílio e suporte.

À FAPESP pelo financiamento da minha bolsa de estudos.

A todas as pessoas que contribuíram para este trabalho, direta ou indiretamente, conhecidos, amigos e familiares.

## RESUMO

As cisteíno-peptidases (CPs) são enzimas proteolíticas que possuem um resíduo de cisteína em seu sítio ativo. As legumaínas de plantas são CPs conhecidas como enzimas de processamento vacuolar (VPE) e participam nos processos de maturação de sementes, germinação, senescência, resposta a estresses, morte celular programada no desenvolvimento da planta ou em resposta ao ataque de patógenos. Embora existam diversos estudos com legumaínas de plantas, a maioria deles está relacionada às funções das legumaínas em sementes de dicotiledôneas. Até o momento, apenas uma legumaína de cana-de-açúcar havia sido descrita. Neste presente trabalho, foi realizada a caracterização de uma nova legumaína de cana-de-açúcar, denominada CaneLEG2. A CaneLEG2 recombinante foi produzida em sistema de expressão heteróloga *Pichia pastoris* e sua caracterização cinética mostrou que ela apresenta auto-ativação e atividade em pH ácido, características comuns em legumaínas de plantas. Esse estudo ainda demonstrou que a cistatina de cana-de-açúcar CaneCPI-3 exerce uma forte inibição sobre a atividade da CaneLEG2, sugerindo que essa cistatina pode atuar na regulação de cisteíno-peptidases endógenas. Os resultados obtidos neste trabalho ajudarão no entendimento das funções exercidas pela CaneLEG2 e CaneCPI-3 na cana-de-açúcar.

**Palavras-chave:** cisteíno-peptidase, cistatina, legumaína, *Pichia pastoris*, *Saccharum sp.*



## ABSTRACT

Cysteine proteases (CPs) are proteolytic enzymes which have a cysteine residue at its active site. Plant legumains are CPs known as vacuolar processing enzymes and they play key roles in seed maturation, germination, senescence, stress response, programmed cell death during development and defense against pathogens. Although there are many studies about plant legumains, most of them is related to legumain functions in seeds of dicotyledonous. To date, only one legumain from sugarcane has been described. In this study, it was performed the characterization of a new sugarcane legumain, named CaneLEG2. The recombinant CaneLEG2 was produced in the heterologous expression system *Pichia pastoris* and its kinetic characterization showed that it exhibits self-activation and activity under acidic pH, which are common features of plant legumains. This study also demonstrated that the sugarcane cystatin CaneCPI-3 has a strong inhibition over the CaneLEG2 activity, suggesting that this cystatin may participate of the regulation of endogenous cysteine proteases. The results obtained in this work will support the understanding of the functions of CaneLEG2 and CaneCPI-3 in sugarcane.

**Key-words:** cysteine protease, cystatin, legumain, *Pichia pastoris*, *Saccharum sp.*

## LISTA DE FIGURAS

<b>Figura 1.1.</b> Estrutura tridimensional da caspase-1 (família C14).....	3
<b>Figura 1.2.</b> Estrutura das legumainas em plantas e local de remoção dos pró-peptídeos.....	4
<b>Figura 1.3.</b> Modelo hipotético para o sistema de processamento vacuolar em plantas.....	7
<b>Figura 1.4.</b> Interação entre cistatina e a cisteíno-peptidase alvo.....	9
<b>Figura 3.1.</b> Mapa do vetor de expressão pGAPZaA (Invitrogen) e sequência do sítio de múltipla clonagem (MCS) .....	15
<b>Figura 4.1.</b> Gráfico obtido no programa Blast2GO mostrando a distribuição de espécies com maior similaridade encontradas na análise.....	24
<b>Figura 4.2.</b> Árvore filogenética construída para clusterização das sequências de legumainas buscadas no SUCEST.....	26
<b>Figura 4.3.</b> Sequência de nucleotídeos e proteína predita da CaneLEG2.....	27
<b>Figura 4.4.</b> Alinhamento entre CaneLEG2 de <i>Saccharum</i> sp e outras legumainas feito no programa Multialin.....	28
<b>Figura 4.5.</b> Provável sítio de glicosilação da CaneLEG2.....	29
<b>Figura 4.6.</b> Amplificação da ORF da legumaina de cana-de-açúcar (CaneLEG2) a partir do clone SCJFRT2053G04, utilizando <i>primers</i> específicos.....	30
<b>Figura 4.7.</b> PCR de colônia para identificação de transformantes de <i>E. coli</i> DH5 com o plasmídeo recombinante.....	31
<b>Figura 4.8.</b> Análise de restrição do plasmídeo recombinante.....	31
<b>Figura 4.9.</b> Plasmídeo recombinante pGAPZ A-CaneLEG2 linearizado.....	32
<b>Figura 4.10.</b> PCR das colônias de KM71H transformadas com plasmídeo recombinante pPGAPZ A.....	33
<b>Figura 4.11.</b> Expressão da CaneLEG2 recombinante em <i>P. pastoris</i> .....	34
<b>Figura 4.12.</b> Expressão heteróloga e purificação da CaneLEG2 junto com o inibidor CaneCPI-3CT.....	35
<b>Figura 4.13.</b> Detecção da CaneLEG2 recombinante produzida em <i>P. pastoris</i> .....	36
<b>Figura 4.14.</b> Purificação e separação da CaneLEG2 e o inibidor CaneCPI-3CT.....	37
<b>Figura 4.15.</b> Caracterização cinética da atividade enzimática da CaneLEG2 recombinante...38	
<b>Figura 4.16.</b> Inibição da CaneLEG2 por CaneCPI-3CT, CaneCPI-3, CaneCPI-3DeIC e Ac-YVAD-CHO .....	41

## LISTA DE TABELAS

<b>Tabela 3.1.</b> Oligonucleotídeos utilizados para clonagem da CaneLEG2.....	14
<b>Tabela 4.1.</b> Descrição das bibliotecas que compõem o SUCEST.....	23
<b>Tabela 4.2.</b> Valores das constantes de inibição ( $K_i$ ) de diferentes inibidores contra a CaneLEG2.....	39

## LISTA DE ABREVIATURAS E SIGLAS

$\mu\text{F}$	microfaraday
$\mu\text{g}$	microgramas
$\mu\text{L}$	microlitros
$\mu\text{M}$	micromolar
	ohm
$^{\circ}\text{C}$	graus Celsius
	comprimento de onda
aa	aminoácidos
Ac-YVAD-CHO	Acetyl-Tyr-Val-Ala-Asp-aldehyde
AOX	álcool oxidase
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
cDNA	DNA complementar
cm	centímetros
CP	cisteíno-peptidase
CPI	cysteine protease inhibitor
DNA	ácidos desoxirribonucléico
dNTP	desoxirribonucleotídeos trifosfatos
DO	densidade óptica
DTT	ditiotreitól
E-64	transepoxi-succinil-L-leucilamido-(4-guanidino)butano
E.C.	Enzyme Commission
EST	<i>expressed sequence tag</i>
g	gramas
<i>g</i>	gravidade
GAPDH	glicerol-3-fosfato-dehidrogenase
GH	glicosil-hidrolase
h	hora
HCl	ácido clorídrico
kb	quilobases
kDa	quilodalton
kV	quilovolts
<i>K<sub>i</sub></i>	constante de inibição
<i>K<sub>m</sub></i>	constante de Michaelis-Menten

L litro  
LB Luria-Bertani  
M molar  
MCP morte celular programada  
mg miligrama  
MgCl<sub>2</sub> cloreto de magnésio  
min minuto  
mL mililitros  
mM milimolar  
MW *molecular weight* (massa molecular)  
Na<sub>2</sub>HPO<sub>4</sub> fosfato de sódio dibásico  
NaHPO<sub>4</sub> fosfato de sódio monobásico  
NaCl cloreto de sódio  
ng nanogramas  
nm nanômetros  
ORF fase aberta de leitura (*open reading frame*)  
pb pares de bases  
PCR reação em cadeia da polimerase (*polymerase chain reaction*)  
pH potencial de hidrogênio  
PVDF polivinilideno difluorido  
rpm rotações por minuto  
s segundos  
SDS-PAGE eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio  
SUCEST Sugarcane Expressed Tags Project  
U unidade enzimática  
UAF unidade arbitrária de fluorescência  
UV ultra violeta  
V volts  
VS versus  
Z-A-A-N-MCA carbobenzoxicarbonil-Ala-Ala-Asn-7-amido-4-metilcoumarina  
Z-V-A-N-MCA carbobenzoxicarbonil-Val-Ala-Asn-7-amido-4-metilcoumarina  
YPD meio dextrose, peptona e extrato de levedura (*yeast extract peptone dextrose*)  
YPDS meio sorbitol, dextrose, peptona, extrato de levedura (*yeast extract peptone dextrose sorbitol*)

## SUMÁRIO

<b>AGRADECIMENTOS</b> .....	i
<b>RESUMO</b> .....	ii
<b>ABSTRACT</b> .....	iii
<b>LISTA DE FIGURAS</b> .....	iv
<b>LISTA DE TABELAS</b> .....	v
<b>LISTA DE ABREVIATURAS E SIGLAS</b> .....	vi
<b>1. INTRODUÇÃO</b> .....	1
1.1. Peptidases.....	1
1.1.1. Legumainas.....	2
1.2. Cistatinas.....	8
<b>2. OBJETIVOS</b> .....	12
<b>3. MATERIAL E MÉTODOS</b> .....	13
3.1. Análises <i>in silico</i> .....	13
3.2. Clonagem da legumaina no vetor de expressão heteróloga.....	14
3.3. Expressão heteróloga em levedura <i>Pichia pastoris</i> .....	17
3.4. Produção e purificação da proteína recombinante.....	19
3.5. Ensaio de imunodeteção.....	21
3.6. Caracterização cinética da CaneLEG2.....	21
3.6.1. Cálculo da constante de Michaelis Menten ( <i>K<sub>m</sub></i> ).....	21
3.6.2. Ensaio de inibição enzimática.....	22
<b>4. RESULTADOS E DISCUSSÃO</b> .....	23
4.1. Análises <i>in silico</i> .....	23
4.2. Clonagem da CaneLEG2 no vetor de expressão pGAPZ A.....	30
4.3. Transformação de células competentes de <i>P. pastoris</i> e expressão da CaneLEG2 recombinante.....	33
4.4. Indução da expressão e purificação da CaneLEG2 recombinante.....	34
4.5. Ensaio de imunodeteção.....	36
4.6. Separação do complexo enzima-inibidor e ativação da CaneLEG2.....	36
4.7. Determinação da constante de Michaelis Menten da enzima CaneLEG2.....	37
4.8. Ensaio de inibição da atividade da CaneLEG2 <i>in vitro</i> .....	39
<b>5. CONCLUSÃO</b> .....	44
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	45

# 1. INTRODUÇÃO

## 1.1. Peptidases

As peptidases são hidrolases que atuam na hidrólise de ligações peptídicas (EC 3.4). Estas enzimas são amplamente distribuídas entre os organismos vivos e representam cerca de 2% de todo produto gênico (IUMBM, 1992; RAWLINGS *et al.*, 2014). As peptidases são divididas em exopeptidases e endopeptidases. As exopeptidases removem de um a três aminoácidos na extremidade amino (aminopeptidases) ou carboxi-terminal (carboxipeptidases) de um peptídeo. Já as endopeptidases, hidrolisam ligações presentes no interior da cadeia polipeptídica (RAWLINGS; MORTON; BARRETT, 2007).

As peptidases diferem quanto à sua localização celular, função fisiológica, especificidade de substrato e inibidores. O estudo da estrutura do sítio ativo das peptidases indicou que este se localiza em um sulco na superfície da proteína entre domínios estruturais próximos. A especificidade de um substrato está relacionada à capacidade de seus aminoácidos ligarem-se aos sítios dispostos ao longo do sulco, em um ou ambos os lados, do sítio catalítico da enzima responsável pela hidrólise do substrato. Por isso, foi proposto um modelo de especificidade peptidase-substrato em que subsítios específicos (chamados S) presentes no sítio catalítico da enzima acomodam os resíduos de aminoácidos do substrato. Os subsítios são numerados do sítio catalítico da enzima para a região amino (S1, S2, S3, Sn) ou carboxi-terminal (S1', S2', S3', Sn') do substrato. Os aminoácidos do substrato que interagem com cada subsítio são designados de forma similar (P1, P2, P3, Pn ou P1', P2', P3', Pn') (SCHECHTER; BERGER, 1997).

As peptidases são divididas em seis classes principais, de acordo com os resíduos de aminoácidos catalíticos presentes no sítio ativo. As principais classes são: serino-peptidases (EC 3.4.21) que possuem no centro ativo um resíduo de serina envolvido no processo catalítico, as treonino-peptidases (EC 3.4.25) que possuem como resíduo catalítico a treonina, as aspartil-peptidases (EC 3.4.23), cuja atividade catalítica depende de um ou dois resíduos de ácido aspártico; as metalo-peptidases (EC 3.4.24) utilizam um íon metálico, geralmente o  $Zn^{2+}$ , em seu mecanismo de ação catalítica; as glutamil-

peptidases que utilizam a díade catalítica glutamato/glutamina e as cisteíno-peptidases (EC 3.4.22) que possuem um resíduo de cisteína em seu sítio ativo (RAWLINGS *et al.*, 2014).

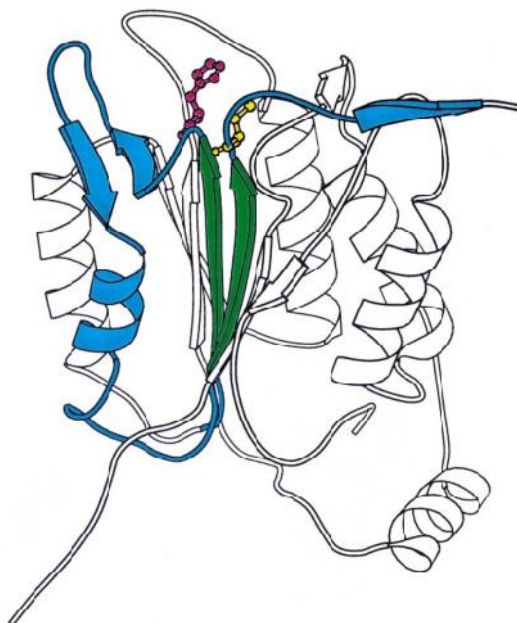
As classes de endopeptidases ainda são agrupadas de acordo com suas similaridades estruturais. Acredita-se que esta classificação possa refletir as características evolutivas dessas proteínas. Portanto, as endopeptidases podem ser agrupadas em famílias e clãs. As famílias são formadas por grupos de enzimas que possuem relação evolutiva com pelo menos outra proteína da mesma família, sendo que essa relação pode ser na sequência completa ou mesmo na região da sequência responsável pela atividade catalítica. Os clãs agrupam famílias que, apesar da ausência de similaridade significativa na sequência, apresentam indícios de relação evolutiva como, por exemplo, a ordem linear dos resíduos catalíticos na estrutura terciária da proteína (RAWLINGS *et al.*, 2014).

### 1.1.1 Legumainas

As enzimas proteolíticas em plantas atuam em diversos processos biológicos. Suas principais funções incluem degradação, maturação de proteínas, senescência, formação de xilema, embriogênese, eventos relacionados à morte celular programada e resposta a estresse biótico e abiótico (MOSOLOV; VALUEVA, 2006). A classe das cisteíno-peptidases (CP) é composta por 105 famílias agrupadas em 10 clãs. As legumainas fazem parte da família C13 das CP. Esta família, também conhecida como família da legumaina, pertence ao clã CD de CP (RAWLINGS *et al.*, 2014). As enzimas desta família foram primeiramente isoladas de sementes maduras de *Ricinus communis*, de cotilédones de *Glycine Max* e de sementes maduras de *Canavalia ensiformis* (HARA-NISHIMURA; INOUE; NISHIMURA, 1991; SCOTT *et al.*, 1992; ABE *et al.*, 1993). Essas enzimas são caracterizadas por clivar especificamente ligações peptídicas no lado C-terminal de um resíduo de asparagina, ou ácido aspártico, sendo também denominadas asparaginil endopeptidases (HARA-NISHIMURA *et al.*, 2005). Foram relatadas semelhanças entre as enzimas da família C13 e as da família das caspases (C14) tanto na sequência, como na estrutura secundária ao redor dos resíduos



do sítio ativo, indicando que pode haver um enovelamento similar entre as enzimas dessas famílias (Figura 1.1).



**Figura 1.1.** Estrutura tridimensional da caspase-1 humana (família C14) (P42573), mostrando em verde as duas fitas que suportam os resíduos catalíticos histidina (rosa) e cisteína (amarelo). A região em azul foi utilizada para comparação da sequência e estrutura secundária entre as enzimas da família C13 e C14. CHEN *et al.*, 1998.

As peptidases da família C13 possuem os resíduos catalíticos cisteína e histidina, dispostos na ordem His-Cys, precedidos por um bloco de quatro aminoácidos hidrofóbicos presentes nas fitas que suportam os resíduos catalíticos cisteína e histidina, comum às CP (CHEN *et al.*, 1998; RAWLINGS *et al.*, 2008). O mecanismo catalítico da legumaína humana descrito por Dall e Brandstetter (2013) envolve a ligação do oxigênio do grupo carbonila do resíduo P1 a um bolso de oxianion trivalente formado pelos resíduos catalíticos cisteína (Cys189) e histidina (His148), com a participação de uma glicina (Gly149).

As legumaínas são sintetizadas como um polipeptídeo precursor contendo um peptídeo sinal, o qual é removido no retículo endoplasmático rugoso (RER), e dois pró-peptídeos, um na região N-terminal e outro na região C-terminal (MÜNTZ; SHUTOV, 2002). A pró-legumaína é transportada para o vacúolo na sua forma inativa e sua ativação depende da remoção dos pró-peptídeos, ou pelo menos do pró-peptídeo C-terminal, como observado em *Ricinus communis* (Figura 1.2). Ao chegar ao seu sítio de destino ocorre o

processamento da pró-legumaina de forma autocatalítica e dependente do pH do meio, uma vez que os sítios de clivagem dos pró-peptídeos são flanqueados por resíduos de asparagina ou ácido aspártico (HIRAIWA; NISHIMURA; HARA-NISHIMURA, 1999; KUROYANAGI; NISHIMURA; HARA-NISHIMURA, 2002; HATSUGAI *et al.*, 2006; SANTOS-SILVA *et al.*, 2012).



**Figura 1.2.** Estrutura das legumainas em plantas e local de remoção dos pró-peptídeos. SP: peptídeo sinal; PPN: pró-peptídeo N-terminal. PPC: pró-peptídeo C-terminal. A proteína madura está representada em azul. RE: retículo endoplasmático; VA: vacúolo. SANTOS-SILVA, L. K., 2012, adaptado de MÜNTZ; SHUTOV, 2002.

Em estudos realizados por Zhao *et al.* (2014) foi feita uma comparação da estrutura cristalizada de uma asparaginil endopeptidase (AEP) de camundongos (*M. musculus*) processada sob diferentes condições de pH. Os resultados mostraram que quando incubada em pH 3,5 e cristalizada em pH 6,7, a estrutura da AEP apresentava apenas o domínio nuclear (*core domain*) sem a presença do domínio C-terminal (CTD). Quando a AEP era incubada em pH 7,5 e cristalizada em pH 7,0, sua estrutura era claramente dividida em duas partes: o domínio nuclear com fitas e  $\alpha$ -hélices e um domínio *cap* completamente formado por  $\alpha$ -hélices localizado exatamente acima da região conservada que contém os resíduos catalíticos histidina e cisteína, cobrindo o bolso de ligação ao substrato. Esses estudos explicam a ativação da legumaina em pH ácido e mostram ainda que a AEP parcialmente ativada pode retornar à sua forma de pró-enzima inativa quando o pH é elevado para 7,5.

Em plantas, as legumainas também são conhecidas como enzimas de processamento vacuolar (*Vacuolar Processing Enzyme*, VPE). A maioria das VPEs, após sua síntese, é direcionada aos vacúolos, com exceção da nucelaína, uma VPE de cevada, cuja presença foi detectada na parede celular do nucelo (LINNESTAD *et al.*, 1998, MÜNTZ *et al.*, 2002). Os vacúolos são organelas multifuncionais que ocupam grande parte do volume celular e atuam na manutenção da pressão de turgor, homeostase do protoplasma, armazenamento de produtos metabólicos, captação de xenobióticos e digestão

de componentes celulares. O vacúolo ainda possui papel na defesa contra micro-organismos e insetos herbívoros. Os vacúolos podem ser classificados em dois diferentes tipos. Os vacúolos de armazenamento de proteínas (VAP) possuem grandes quantidades de proteínas, entre elas proteínas de defesa e de reserva que servem como fonte de carbono, nitrogênio e enxofre para a germinação da semente e crescimento da planta jovem. Já os vacúolos líticos possuem enzimas hidrolíticas para a degradação de material celular, que acontece, por exemplo, durante o processo de morte celular programada (HERMAN; LARKINS, 1999; HARA-NISHIMURA; HATSUGAI, 2011).

A morte celular programada é um processo de suicídio celular muito bem regulado para crescimento e sobrevivência em eucariotos. Em plantas, MCP ocorre no desenvolvimento, bem como durante xilogênese, embriogênese, formação de aerênquima, diversos processos reprodutivos, desenvolvimento de sementes e senescência das folhas. Além do seu papel no desenvolvimento de plantas, o suicídio celular contribui amplamente na defesa contra estresses ambientais como ozônio e radiação ultravioleta, e na defesa contra o ataque de patógenos (HATSUGAI *et al.*, 2006). Proteases também possuem um papel importante na imunidade inata de plantas. Estudos recentes mostraram o envolvimento de VPE na transdução de sinal que induz o fechamento de estômatos e respostas de defesa, incluindo a expressão de genes de defesa e resposta hipersensitiva (ZHANG; ZHENG; ZHANG *et al.*, 2010).

A MCP em células animais é geralmente executada por caspases. Essas pertencem a uma classe de cisteíno proteases que mostram um alto grau de especificidade, com uma exigência absoluta por clivagem adjacente a um resíduo de aspartato e uma sequência de reconhecimento composta de pelo menos quatro aminoácidos N-terminais para este sítio de clivagem. Em células animais, um número de características morfológicas e bioquímicas relacionadas à morte celular (por exemplo, condensação da cromatina e fragmentação nuclear e de DNA) é atribuído ao processamento de moléculas-alvo específicas mediado por caspases. Essas características morfológicas e bioquímicas são também observadas em células vegetais em processo de MCP. Isso implica numa similaridade no mecanismo de MCP em animais e plantas e sugere que proteases com preferência pelos mesmos substratos da caspase podem estar envolvidas. A morte celular programada em plantas pode

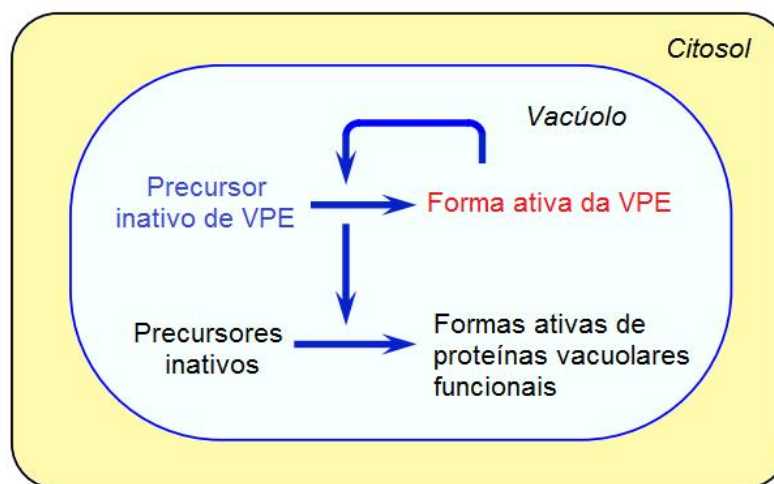
ser efetivamente bloqueada por inibidores de caspase humana sintéticos e macromoleculares. Extratos de plantas têm mostrado atividade de clivagem em substratos sintéticos fluorogênicos e naturais de caspase (WOLTERING, 2010).

As VPEs estão envolvidas nos processos de degradação e mobilização de proteínas de reserva durante a germinação da semente e desenvolvimento da planta jovem, inclusive atuando no processamento de pró-CPs do tipo papaína envolvidas na degradação das proteínas de reserva (OKAMOTO; MINAMIKAWA, 1999).

A análise da expressão das VPEs em *Arabidopsis* levou à classificação destas em vegetativas ou de sementes, de acordo com os tecidos onde houve maior expressão. Apesar da expressão de VPEs vegetativas ter sido observada em sementes, a classificação das VPEs ainda continua a ser usada baseada nas funções que as enzimas exercem na planta (KINOSHITA *et al.*, 1999; MÜNTZ, BLATTNER; SHUTOV, 2002; GRUIS; SCHULZE; JUNG, 2004). As enzimas de processamento vacuolar vegetativas estão envolvidas majoritariamente nos processos de senescência, resposta a estresse e morte celular programada (MCP) durante o desenvolvimento da planta ou resposta a patógenos (HATSUGAI *et al.*, 2004; ROJO *et al.*, 2004). As VPEs de sementes estão relacionadas ao processamento e mobilização de proteínas de reserva e no processo de MCP durante o desenvolvimento da semente (SHIMADA *et al.*, 2003; NAKAUNE *et al.*, 2005). Análises da expressão dos promotores de VPEs de *Arabidopsis* mostraram que a VPE, VPE da semente, foi mais expressa em sementes secas, eixo embrionário e cotilédones, enquanto as VPEs vegetativas (VPE e VPE) foram expressas principalmente em tecidos durante o processo de MCP, raízes e folhas senescentes (KINHOSHITA *et al.*, 1999). A expressão de uma VPE de semente, VPE, foi observada em sementes em desenvolvimento (GRUIS, *et al.*, 2002; NAKAUNE *et al.*, 2005). Uma mutação no gene da enzima VPE em *Arabidopsis* levou a acúmulo de precursores de proteínas de reserva, globulinas 12S e albuminas 2S. Mesmo com a atividade compensatória das VPEs VPE e VPE, a maior parte da atividade enzimática nas sementes era resultado da atuação da VPE (SHIMADA *et al.*, 2003). Chen *et al.*, (2004) demonstraram que houve um aumento na expressão da VPE quando folhas de batata-doce (*Ipomoea batatas*) eram tratadas com indutores de senescência. Isso ocorre porque durante a senescência das folhas

há degradação das proteínas por peptidases e reutilização das mesmas em partes não-senescentes da planta. Ademais, Rojo *et al.* (2003) demonstraram que há a degradação de uma invertase vacuolar em folhas senescentes resultante da ação da VPE, pois plantas mutantes VPE apresentaram os mesmos níveis de invertase mesmo após o início da senescência.

Algumas proteínas de reserva das plantas são sintetizadas como precursores que são processados na sua forma madura assim que se acumulam no vacúolo. Esse acúmulo de proteínas nos VAP geralmente ocorre no desenvolvimento das sementes e, durante a germinação, elas servirão como fonte de nutrientes para o crescimento da planta (HIGGINS, 1984; KEMBHAVI *et al.*, 1993). Jung *et al.* (1998) demonstraram que as globulinas, proteínas de reserva, são sintetizadas como precursores no retículo endoplasmático rugoso. Após sua síntese e clivagem do peptídeo sinal, elas são organizadas em trímeros que são transportados aos VAPs e processados por asparaginil endopeptidases, mostrando que as VPEs são responsáveis pela maturação de precursores dessas proteínas de reserva (Figura 1.3).



**Figura 1.3.** Modelo hipotético para o sistema de processamento vacuolar em plantas. Uma variedade de proteínas vacuolares é sintetizada no retículo endoplasmático como pró-proteínas precursoras e então são transportadas aos vacúolos. Elas então são proteoliticamente processadas em suas formas maduras pelas VPEs. A própria VPE é autocataliticamente convertida na sua forma madura, funcionando como enzima chave no sistema de processamento vacuolar. Figura adaptada de Hatsugai *et al.*, 2006.

A cana-de-açúcar tem sido alvo de estudos devido à sua importância no cenário econômico nacional e internacional. Pesquisas têm sido realizadas

buscando a obtenção de variedades mais produtivas e com menos perdas na produção, tanto por fatores bióticos como abióticos. Portanto o estudo do perfil da expressão gênica durante o desenvolvimento da planta e em condições de estresse é importante para a compreensão dos seus processos fisiológicos.

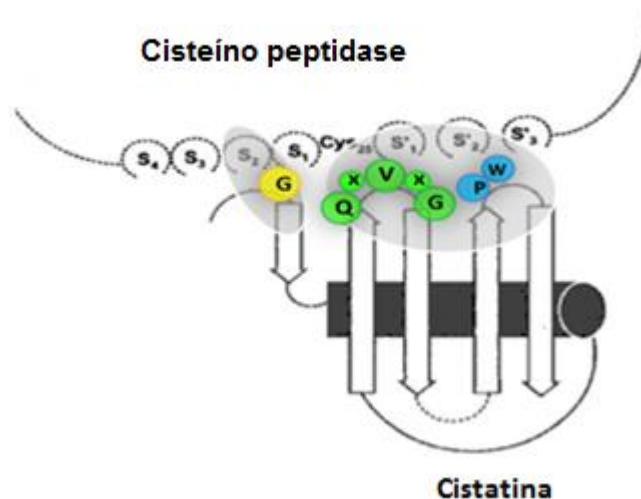
Estudos realizados em nosso laboratório, por Santos-Silva *et al.* (2012) com uma legumaina de cana-de-açúcar (CaneLEG) revelaram que esta enzima possui as características principais de VPEs, como auto-ativação e atividade em pH ácido. Estes estudos também mostraram que a atividade da CaneLEG é fortemente inibida quando incubada com a cistatina 3 da cana-de-açúcar (CaneCPI-3). Análises quantitativas da expressão dos genes de CaneLEG e CaneCPI-3 indicaram um padrão de expressão tecido-específico durante o crescimento da cana, com forte acúmulo de transcritos de CaneLEG durante o desenvolvimento dos entrenós. Além disso, a expressão de genes envolvidos em determinados processos fisiológicos também pode ser induzida ou reprimida pela ação de fitohormônios. Santos-Silva *et al.* (2012) demonstraram que a expressão de uma legumaina de cana foi fortemente induzida por tratamento com ácido abscísico (ABA), sugerindo que estes genes podem estar envolvidos em resposta a estresse relacionado a ABA.

## **1.2 Cistatinas: Inibidores naturais de cisteíno-peptidases**

As cistatinas são inibidores reversíveis de cisteíno peptidases e seu mecanismo de ação baseia-se na inibição competitiva, bloqueando o acesso do substrato específico ao sítio ativo da enzima (TURK; BODE, 1991; BENCHABANE *et al.*, 2010). Estudos descrevem as cistatinas como inibidores de CP da família da papaína (TURK; BODE, 1991; ARAI *et al.*, 2002; OLIVA *et al.*, 2004). Entretanto, enzimas da família da legumaina também podem ser inibidas por algumas cistatinas, num sítio reativo diferente do utilizado para inibição de CP do tipo papaína (ALVAREZ-FERNANDEZ *et al.*, 1999; CHENG *et al.*, 2006).

A maioria das cistatinas possui em sua composição uma folha dobrada, contendo cinco fitas antiparalelas, organizada em volta de uma hélice de cinco voltas. A inibição das enzimas da família da papaína se dá por meio da interação entre três regiões conservadas das cistatinas e a cisteíno-

peptidase alvo: uma sequência N-terminal contendo um resíduo de glicina que interage com o subsítio S2 da enzima; um motivo Gln-X-Val-X-Gly altamente conservado na região central formando o *loop* 1 e uma região contendo um resíduo de triptofano, localizado no *loop* 2 que apresentam tamanhos e formas que se encaixam nos subsítios S1' e S2' do sítio ativo da enzima (TURK; BODE, 1991; BENCHABANE *et al.*, 2010) (Figura 1.4).



**Figura 1.4.** Interação entre cistatina e cisteína peptidase alvo, em destaque as três regiões da cistatina envolvidas na inibição: em amarelo o segmento N-terminal, contendo o resíduo de glicina (G) conservado, que interage com a peptidase no subsítio S2; em verde a alça central, contendo o motivo Q-X-V-X-G, e em azul a região C-terminal, contendo o resíduo conservado de triptofano (W), que constituem as outras regiões de contato com a peptidase em S1' e S2', respectivamente. Figura adaptada de ALVAREZ-FERNANDEZ e ABRAHAMSON, 2006.

Como há um grande número de cistatinas descritas, elas foram organizadas em uma superfamília subdividida, inicialmente, em três famílias de cistatinas (BARRETT *et al.*, 1986). A inclusão de uma nova família contendo as cistatinas de plantas, ou fitocistatinas, foi sugerida por Kondo *et al.* (1991). As fitocistatinas possuem algumas características semelhantes à família das estefinas, como ausência de ligações dissulfeto e resíduos de cisteína. Também assemelha-se à família das cistatinas, como o tamanho e a similaridade na sequência primária. Além disso, as fitocistatinas possuem exclusivamente a sequência consenso L-A-R-[FY]-A-[VI]-X(3)-N, formando uma hélice na região N-terminal. A análise filogenética de cistatinas animais e de plantas mostrou o agrupamento das fitocistatinas em um único ramo, distinto

das famílias 1 e 2, justificando sua classificação numa família própria (MARGIS; REIS; VILLERET, 1998).

As fitocistatinas estão envolvidas na regulação da atividade das peptidases endógenas durante na maturação de sementes e na defesa contra cisteíno-peptidases exógenas de insetos e nematoides, que geralmente possuem essas peptidases em seus intestinos (ARAI, *et al.*, 2002).

As fitocistatinas foram identificadas em diversas plantas, incluindo mono e dicotiledôneas. A primeira fitocistatina foi isolada de sementes de arroz (ABE *et al.*, 1987). Desde então, já foram caracterizadas cistatinas em plantas como milho (YAMADA *et al.*, 2000), cana-de-açúcar (SOARES-COSTA *et al.*, 2002; GIANOTTI *et al.*, 2006), batata (ANNADANA *et al.*, 2003), feijão (DIOP *et al.*, 2004), trigo (CORRE-MENGUY *et al.*, 2002), entre outras. As fitocistatinas foram inicialmente descritas por sua capacidade inibitória contra enzimas da família da papaína. Entretanto, Martinez *et al.* (2007) descreveram a inibição de enzimas da família da legumaína pela cistatina de cevada com massa molecular de 23 kDa, contendo uma região C-terminal estendida.

A estrutura tridimensional da cistatina de arroz, orizacistatina I, apresenta o enovelamento típico observado para as cistatinas animais. Embora as fitocistatinas contendo a extensão C-terminal apresentem enovelamento de sua região N-terminal similar a orizacistatina I, a estrutura tridimensional de sua região C-terminal estendida ainda não foi determinada (CHU *et al.*, 2011). As funções das fitocistatinas estão relacionadas à regulação de cisteíno-peptidases endógenas durante a maturação de sementes, degradação de proteínas nos órgãos de reserva, crescimento e proliferação celular, senescência e morte celular programada (SOLOMON, *et al.*, 1999, SUGAWARA *et al.*, 2002; HWANG *et al.*, 2009; WEEDA; KUMAR; KNOWLES, 2009; TIAN *et al.*, 2009). Entretanto, pouco se sabe a respeito de quais CP endógenas são alvos desses inibidores. Yamada *et al.* (2000) confirmaram a interação de fitocistatina com CP do tipo papaína *in vivo*, após a purificação de uma cistatina de milho com sua peptidase alvo, sendo necessário o tratamento do complexo com SDS para liberação da fitocistatina e ativação da peptidase (YAMADA *et al.*, 2000).

As fitocistatinas também atuam na defesa da planta em resposta a ferimentos ou ataque por patógenos e insetos. Alguns fitopatógenos e insetos



secretam cisteíno-peptidases para facilitar o processo de patogênese e degradação de proteínas da planta (HABIB; FAZILI, 2007). Além disso, fitocistatinas são capazes de inibir CP digestivas de insetos e o crescimento de microorganismos fitopatogênicos *in vitro* (ZHAO *et al.*, 1996; YANG; YEH, 2005; CARRILLO *et al.*, 2011). Também foi observado que as cistatinas são super-expressas em condições de estresse abiótico como alta salinidade, seca, oxidação e frio (PERNAS; SANCHEZ-MONGE; SALCEDO, 2000; DIOP *et al.*, 2004). Portanto, o uso de fitocistatinas para produção de plantas transgênicas mais resistentes a patógenos, insetos e estresse abiótico vem se tornando cada vez mais crescente (URWIN; GREEN; ATKINSON, 2003; HAQ; ATIF; KHAN, 2004; HABIB; FAZILI, 2007; ZHANG *et al.*, 2008) e os estudos do perfil inibitórios desses inibidores é importante para a caracterização inicial e entendimento da função das fitocistatinas na planta.

No laboratório de Biologia Molecular (LBM – DGE – UFSCar) foi realizada a caracterização da primeira cistatina de cana-de-açúcar, denominada CaneCPI-1, além de outras três cistatinas da mesma planta. Essas cistatinas são capazes de inibir de forma eficiente catepsinas humanas, assim como outras CP (SOARES-COSTA *et al.*, 2002; OLIVA *et al.*, 2004; GIANOTTI *et al.*, 2006). Estudos realizados com a CaneCPI-1 mostraram que esse inibidor apresentou atividade antifúngica contra *Trichoderma reesei*, indicando que esta cistatina pode estar envolvida na defesa da planta contra patógenos. Foi demonstrado também em nosso laboratório que a CaneCPI-3, uma cistatina de cana-de-açúcar com a região C-terminal estendida, possui uma forte capacidade de inibição de VPEs, com a asparagina do motivo SNSL na extensão C-terminal sendo responsável pela atividade inibitória (SANTOS-SILVA *et al.*, 2012).

## 2. OBJETIVOS

O presente trabalho teve como objetivos gerais encontrar novas legumaínas de cana-de-açúcar por meio de uma busca no banco de dados de EST (SUCEST) e realizar a expressão heteróloga e caracterização de pelo menos uma das enzimas encontradas. Nesse sentido, os objetivos específicos propostos, foram:

- Identificação das sequências de legumaínas presentes no banco de dados do SUCEST;
- Alinhamento das sequências encontradas e clusterização;
- Isolamento e sequenciamento de um clone proveniente do SUCEST contendo a ORF de uma provável legumaína;
- Análise *in silico* da sequência da legumaína;
- Clonagem da ORF da legumaína em vetor específico para expressão em levedura;
- Expressão heteróloga em *Pichia pastoris*;
- Purificação da proteína recombinante;
- Caracterização cinética da enzima recombinante;
- Ensaio de inibição da atividade enzimática da legumaína por inibidores sintéticos e a cistatina de cana-de-açúcar CaneCPI-3.

### 3. MATERIAL E MÉTODOS

#### 3.1 Análises *in silico*

Foi realizada uma busca nas sequências disponíveis no NCBI referentes ao banco de dados do SUCEST (*Sugarcane Expressed Sequence Tag*) através de uma busca por similaridade, utilizando o programa TBLastN, utilizando como modelo a sequência da legumaina de cana-de-açúcar, CaneLEG (GenBank ID: JQ083602), já caracterizada no laboratório (LBM-DGE – UFSCAR). Foram selecionadas as sequências que obtiveram um valor de E-value inferior a  $1e^{-3}$ .

As sequências encontradas foram submetidas à análise no programa Blast2GO (CONESA; GÖTZ *et al.*, 2005). O primeiro passo realizado foi uma busca por similaridade com o blastx contra o banco nr do NCBI, utilizando um e-value mínimo de  $e^{-6}$ . Em seguida, foi realizado o mapeamento relacionando os resultados do blast com termos do banco Gene Ontology (GO) cobrindo três domínios: componente celular (C), função molecular (F) e processo biológico (P). As sequências também foram classificadas segundo a nomenclatura da IUBMB (*International Union of Biochemistry and Molecular Biology*) e com o banco Kegg (Enciclopédia de Genes e Genoma de Kioto). A fim de complementar a anotação funcional as sequências foram analisadas no banco Interpro utilizado o programa InterProScan. Essa ferramenta faz uma busca em diferentes bancos de dados com a finalidade de identificar relações entre as sequências, conseguindo assim inferir funções proteicas.

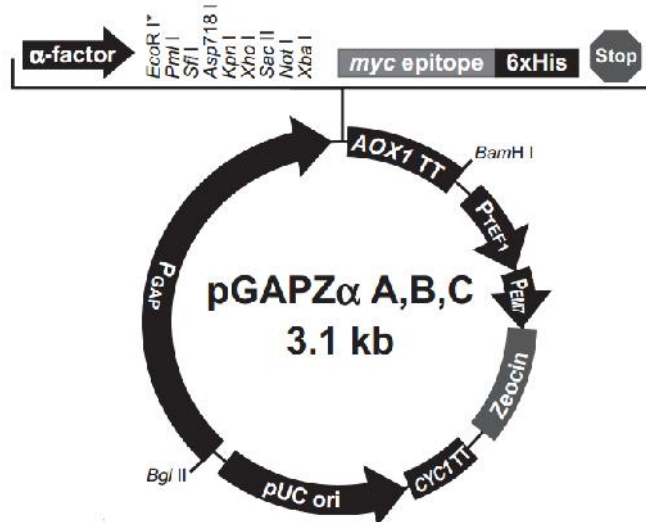
Para verificar se existem tipos diferentes de legumainas na cana-de-açúcar, as sequências foram analisadas no programa MEGA5 (TAMURA, 2011). Foi realizado um alinhamento das sequências encontradas utilizando a ferramenta MUSCLE (Multiple Sequence Comparison by Log-Expectation). As sequências idênticas ou truncadas foram previamente removidas. O método de tratamento de sequências utilizado foi o *Gblocks* com os parâmetros padrões no servidor PhyML 3.0 (GUINDON; GASCUEL, 2003) que elimina regiões mal alinhadas ou divergentes. A árvore filogenética das legumainas foi obtida utilizando o método de máxima verossimilhança. O melhor modelo foi determinado com o critério de informação Akaike (AIC) (AKAIKE, 1974). O suporte estatístico para os nós foi avaliado por *bootstrap* com 1000 pseudoréplicas.

### 3.2. Clonagem da legumaína no vetor de expressão heteróloga

Após obtenção do cDNA da CaneLEG2, foram desenhados oligonucleotídeos contendo sítios de restrição para amplificação por PCR de sua ORF, sem o peptídeo sinal, para clonagem em vetores de expressão. Para amplificação da ORF foram usados oligonucleotídeos *forward* contendo o sítio de restrição para *EcoRI*, e *reverse* contendo o sítio de restrição para a enzima *SalI* (tabela 3.1). O vetor de expressão utilizado para clonagem foi o pGAPZ A (Invitrogen) (Figura 2.1).

**Tabela 3.1.** Oligonucleotídeos utilizados para clonagem da CaneLEG2. Os sítios para as enzimas de restrição *EcoRI* (GAATTC) e *SalI* (GTCGAC) estão sublinhados

Nome	Sequência 5' – 3'
CLeg2EcoRIFw	A <u>AGAATTC</u> GCGGCGGATGGGGCGGAG
CLeg2SalRv	AA <u>AGTCGAC</u> AGCGCTGTGTCCCTGAGC



```

361 GATTATTGGA AACCACCAGA ATCGAATATA AAAGGCGAAC ACCTTTCCCA ATTTTGGTTT
                                     pGAP forward priming site
421 CTCCTGACCC AAAGACTTTA AATTTAATTT ATTTGTCCCT ATTTCAATCA ATTGAACAAC
481 TATTTCGAAA CG ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA
    Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala

532 GCA TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG
    Ala Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr
                                     α-factor signal sequence
583 GCA CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT
    Ala Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp

634 TTC GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG
    Phe Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu

685 TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT
    Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser

Xho I*      Kex2 signal cleavage      EcoR I      Pml I      Sfi I
736 CTC GAG AAA AGA GAG GCT GAA GCT GAATTCAC GTGGCCCA GCCGGCCGTC TCGGATC
    Leu Glu Lys Arg Glu Ala Glu Ala
                                     Ste13 signal cleavage
Asp718 I Kpn I Xho I      Sac II Not I      Xba I      myc epitope
793 GGTACCTCGA GCCGCGGCGG CCGCCAGCTT TCTA GAA CAA AAA CTC ATC TCA GAA GAG
    Glu Gln Lys Leu Ile Ser Glu Glu

                                     polyhistidine tag
851 GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA GTTTTAGC CTTA
    Asp Leu Asn Ser Ala Val Asp His His His His His His ***

905 GACATGACTG TTCCTCAGTT CAAGTTGGGC ACTTACGAGA AGACCGGTCT TGCTAGATTC TAAT

                                     3' AOX1 priming site
969 CAAGAGGATG TCAGAATGCC ATTTGCCCTGA GAGATGCAGG CTCATTTTTT GATACTTTTT TATT

1033 TGTAACCTAT ATAGTATAGG ATTTTTTTTG TCATTTTGGT TCTTCTCG

```

**Figura 2.1.** Mapa do vetor de expressão pGAPZ A (Invitrogen) e sequência do sítio de clonagem múltipla (MCS). No mapa observa-se a região PGAP, referente ao promotor do gene GAPDH; AOX1 TT, que indica o terminador de transcrição do gene AOX1; Zeocin, o gene de resistência ao antibiótico zeocina; -factor, sequência que permite a secreção da proteína; 6xHis, cauda de polihistidina. No MCS está indicado o início e fim da sequência sinal fator, o sinal de clivagem Kex e as enzimas de restrição que podem ser usadas para a clonagem nesse vetor.

A PCR foi realizada num volume final de 25  $\mu$ L contendo 10 ng de DNA, tampão de reação 1X (20 mM Tris-HCl pH 8.8, 10 mM  $(\text{NH}_4)_2\text{SO}_4$ , 10 mM KCl, 1 mg/mL BSA, 1 % (v/v) Triton X-100, 20 mM  $\text{MgSO}_4$ ), 0,2 mM de cada dNTP, 0,4  $\mu$ M de cada oligonucleotídeo e 1,25 U da enzima Pfu DNA polimerase (Thermo). O programa de amplificação consistiu de uma desnaturação inicial de 95 °C por 2 minutos, seguida por 35 ciclos de 95 °C por 30 s, gradientes de 52 °C a 60 °C por 30 s, 72 °C por 3 min e uma extensão final a 72 °C por 10 minutos. Após a reação, os produtos de PCR foram analisados em gel de agarose 1% contendo brometo de etídeo 1  $\mu$ g/mL e visualizados sob luz UV.

O produto de PCR correspondente à CaneLEG2 foi purificado do gel de agarose utilizando o kit *Wizard SV Gel and PCR Clean Up System* (Promega). O material purificado foi digerido com as enzimas de restrição *EcoRI* e *Sall* para obtenção dos fragmentos contendo extremidades coesivas, as quais permitiram a posterior clonagem direcional no vetor de expressão previamente clivado com as mesmas enzimas.

Nas reações de clivagem para obtenção do fragmento da CaneLEG2 e do vetor de expressão contendo extremidades coesivas foram utilizados 1  $\mu$ g de DNA, 10 U *EcoRI*, 10 U *Sall*, tampão Orange 1X (50 mM Tris-HCl pH 7.5, 10 mM  $\text{MgCl}_2$ , 100 mM NaCl, 0,1 mg/mL BSA), em um volume final de 50  $\mu$ L. As reações de clivagem foram mantidas a 37 °C por 16 h. Em seguida, as reações foram analisadas por eletroforese em gel de agarose 1 % e os fragmentos correspondentes à CaneLEG2 e ao pGAPZ A foram recuperados do gel de agarose com o kit *Wizard SV Gel and PCR Clean Up System* (Promega) e utilizados nas reações de ligação. As reações de ligação foram realizadas em volume de 15  $\mu$ L contendo 50 ng de vetor, 130 ng de inserto, tampão de reação 1X (40 mM Tris-HCl pH 7,8, 10 mM  $\text{MgCl}_2$ , 10 mM DTT, 0,5 mM ATP) e 2,5 U de T4 DNA ligase (Fermentas), incubada por 1h a temperatura ambiente.

Um volume de 10  $\mu$ L das reações de ligação foi utilizado para o processo de transformação, por choque térmico, de células de *Escherichia coli* DH5 quimiocompetentes, preparadas pelo método de cloreto de cálcio ( $\text{CaCl}_2$ ) (SAMBROOK; RUSSEL, 2001). As colônias das bactérias transformadas com a construção contendo o gene de interesse foram selecionadas por cultura em meio LB ágar *low salt* (1 % triptona, 0,5 % extrato de levedura, 0,5 % NaCl, pH

7,5) contendo 25 µg/mL de zeocina. Para confirmar a inserção do plasmídeo, foi realizado PCR pela técnica de PCR de colônias, onde 10 colônias foram escolhidas e repicadas com palitos de dente estéreis em tubos de 0,2 mL contendo a mistura para PCR e a amplificação foi realizada como descrito previamente no item 3.2. Os produtos de amplificação foram analisados em gel de agarose 1% corado com brometo de etídeo e visualizados em luz UV.

Após confirmação por PCR, duas colônias positivas foram selecionadas para extração do DNA plasmidial. Essas colônias foram crescidas por 16 h em 10 mL de meio de cultivo LB *low salt* contendo 25 µg/mL de zeocina. O DNA plasmidial foi extraído utilizando o kit *QIAprep Spin Miniprep* (Qiagen) e os plasmídeos recombinantes foram sequenciados para confirmação da ligação do gene em fase correta de leitura com as regiões reguladoras.

### **3.3. Expressão heteróloga em levedura *Pichia pastoris***

Para transformação da linhagem de *P. pastoris* KM71H, o plasmídeo foi linearizado com a enzima de restrição *Xma*JI (Fermentas), dessa forma facilitando a recombinação homóloga com o genoma da levedura. Para linearização foram utilizados 4 µg de plasmídeo contendo o gene em uma reação com volume final de 50 µL, tampão Tango 1X (33 mM Tris-acetato pH 7.9, 10 mM acetato de Mg, 66 mM acetato de K, 0,1 mg/mL BSA), 10 U da enzima *Xma*JI, incubada a 37 °C *overnight*.

Após a linearização foi adicionado 1/10 do volume de acetato de sódio 3 M pH 5,2, 2,5 volumes de etanol 100 % gelado e a solução foi incubada a -80 °C por 15 min e, em seguida, centrifugada a 16000 g por 10 min a 4 °C. O sobrenadante foi descartado e o precipitado lavado com 500 µL de etanol 80 % gelado e centrifugado novamente a 16000 g por 10 min a 4 °C. O sobrenadante foi descartado e após a secagem para evaporação do etanol residual, o material foi ressuspenso em 10 µL de água destilada estéril.

Aproximadamente 1 µg da construção linearizada foi utilizado para transformar células de KM71H eletrocompetentes segundo protocolo descrito por Cregg (2007). Inicialmente, uma colônia de KM71H foi inoculada em 5 mL de meio YPD (1 % extrato de levedura, 2 % peptona, 2 % dextrose) e cultivada a 30 °C, 250 rpm por 9 h. Uma alíquota de 50 µL dessa cultura foi utilizada

para inocular 50 mL de YPD que foi mantido a 30 °C com agitação de 250 rpm até atingir uma DO<sub>600nm</sub> de 1,2. Para uma medida mais precisa da DO<sub>600nm</sub>, a cultura foi diluída 10 vezes em água antes da leitura em espectrofotômetro. A cultura foi centrifugada a 1500 g 4 °C por 5 min e o sobrenadante descartado. Em seguida, as células foram ressuspensas gentilmente em 10 mL de YPD líquido contendo 170 mM de HEPES e, em seguida, foi adicionado 250 µL de ditioneitol (DTT) 1 M, prosseguindo com uma incubação a 30°C por 15 min sem agitação. Em seguida, foi adicionada água destilada gelada (estéril) totalizando um volume final de 50 mL e o material foi centrifugado nas mesmas condições descritas acima. As células coletadas foram lavadas uma vez com 25 mL de água destilada gelada (estéril) e uma vez com 2 mL de sorbitol 1 M e ressuspensas em 100 µL de sorbitol 1 M.

Para transformação de *Pichia*, o DNA linearizado, em volume de até 10 µL, foi misturado com 40 µL de células competentes e, em seguida, transferido para cubetas de 0,2 cm (Bio-Rad) que foram incubadas em gelo por 5 min. A transformação das células competentes foi realizada por eletroporação no aparelho GenePulser II (Bio-Rad) nas seguintes condições: 1500 V, 25 µF, 200 . Imediatamente após o choque, foi adicionado 1 mL de sorbitol 1 M à cubeta e as células foram transferidas para um tubo de 15 mL e mantidas a 30°C por 2h sem agitação. Para seleção dos transformantes 100 e 200 µL de células foram espalhadas em placas de YPDS ágar (1 % extrato de levedura, 2 % peptona, 2 % dextrose, 1 M sorbitol, 2 % ágar) contendo 100 e 500 µg/mL de zeocina e incubadas a 30°C por 3 dias.

Para confirmação da inserção do plasmídeo contendo o gene da CaneLEG2 no genoma da *P. pastoris* os clones obtidos após transformação foram analisados por PCR de colônia conforme descrito por Akada *et al.*, (2000). As colônias transformantes foram repicadas com palito de dente estéril e colocadas em tubos de 1,5 mL contendo 20 µL de SDS 0,2 %, recém-preparado. A solução contendo as células foi fervida por 3 min e centrifugada a 12000 g por 30 s. Um µL do sobrenadante foi utilizado como molde na PCR num volume final de 25 µL contendo 1 % triton X-100, tampão de reação 1X (100 mM Tris-HCl pH 8.8, 500 mM KCl, 1 % Triton X-100, 16 mM MgCl<sub>2</sub>), 200 µM de cada dNTP, 0,4 µM de cada oligonucleotídeo e 1,25 U da enzima Taq DNA polimerase (Sinapse Inc). A amplificação foi realizada em termociclador



T100 (Bio-Rad) utilizando o programa: 94 °C 3 min, 35 ciclos a 94 °C 90 s, 52 °C 90 s, 72 °C 4 min, extensão final 72 °C por 7 min. Após amplificação o material foi analisado em gel de agarose 1% contendo brometo de etídeo 1 µg/mL e visualizado sob luz UV.

Os clones positivos selecionados por PCR foram testados para avaliar as condições de expressão da proteína recombinante. As colônias recombinantes foram avaliadas quanto à produção da enzima CaneLEG2 em um processo de indução de expressão em pequena escala, usando uma placa de cultura de 24 poços (Whatman). As colônias foram inoculadas em 2 mL de meio YPD tamponado com 100 mM de tampão fosfato de potássio, pH 6,0, a 30°C e agitação de 250 rpm. O clone da endoglucanase 3 de *Trichoderma harzianum* foi usado como controle positivo e uma colônia de *P. pastoris* KM71H não transformada foi usada como controle negativo. A cada 24 h, as placas foram centrifugadas a 1500 g por 10 min e foram retiradas alíquotas de 100 µL do sobrenadante, até completar 96 h de indução. Cada um dos intervalos foi analisado em SDS-PAGE 12 % (SAMBROOK; RUSSEL, 2001). O clone selecionado para a produção da enzima recombinante em escala maior foi o que apresentou maior produção de enzima em um menor intervalo de tempo.

#### **3.4. Produção e purificação da proteína recombinante**

Para a produção da CaneLEG2 em maior escala, uma pequena quantidade de massa celular foi inoculada em 10 mL de meio YPD tamponado com 100 mM de tampão fosfato de potássio pH 6,0, e cultivada a 30 °C sob agitação de 250 rpm por 16 h. Posteriormente, foram utilizados 200 µL deste pré-inóculo para inocular 100 mL de meio YPD tamponado com 100 mM de tampão fosfato de potássio pH 6,0. A esse inóculo foi adicionado 1 mg do inibidor CaneCPI-3CT. A indução foi mantida por 48 h, a 30 °C, sob agitação de 250 rpm. Após a indução, a cultura foi centrifugada a 20 °C por 5 min a 3000 g e o sobrenadante filtrado em membrana de nitrocelulose 0,45 µm (Millipore) e purificado por cromatografia de afinidade utilizando uma coluna carregada com 3 mL de resina Ni-NTA Superflow (Qiagen). À coluna pré-equilibrada com 5 volumes de tampão de lise (10 mM de Tris, 50 mM de Na<sub>2</sub>HPO<sub>4</sub>

monohidratado, 100 mM de NaCl, pH 8,0), foi adicionado o sobrenadante da cultura induzida, previamente filtrado. Após ligação da proteína à resina, foi realizada uma lavagem da coluna com 3 volumes de tampão de lise seguindo para eluição da proteína em tampão de lise contendo concentrações crescentes de imidazol (10, 25, 50, 75, 100 e 250 mM), sendo 2 volumes das soluções com 10 a 100 mM e 3 volumes da solução com 250 mM de imidazol. As amostras eluídas foram analisadas em SDS-PAGE 12 % corado com azul de comassie. A enzima CaneLEG2 purificada juntamente com o inibidor CaneCPI-3CT foi dialisada em membrana de diálise 14.000 MW (Viskase) contra 2 L de tampão de lise pH 8,0 com 1 M de NaCl. A diálise foi mantida a 4°C com 2 trocas sucessivas de tampão a cada 1 h de incubação. A alta concentração salina utilizada resultou na dissociação do complexo enzima-inibidor e, portanto, foi necessária uma nova purificação para separar as frações contendo apenas a enzima. As frações dialisadas contendo a enzima CaneLEG2 com o inibidor CaneCPI-3CT foram então passadas por uma coluna de gel filtração, Superdex™ 75 10/300 GL (GE Healthcare) utilizando o sistema de purificação de proteínas Akta FPLC (Amersham Pharmacia Biotech). Após a obtenção da CaneLEG2 purificada e separada do inibidor, as frações que continham a enzima foram dialisadas em membrana de diálise 14000 MW (Viskase) contra 1 L de tampão citrato-fosfato (138,5 mM Na<sub>2</sub>HPO<sub>4</sub>, 30,7 mM ácido cítrico pH 6,4) a 4 °C com uma troca de tampão após duas horas de incubação e a diálise foi mantida por 16 h.

### **3.5. Ensaio de imunodeteção**

As amostras analisadas por *Western Blotting* foram previamente separadas em SDS-PAGE 12 % e transferidas para uma membrana de PVDF (Pierce) com tampão de transferência (20 mM Tris, 55 mM glicina, 20% metanol) durante 2 h, a 150 mA, utilizando o sistema Mini Trans-Blot (BIO-RAD). A membrana foi bloqueada por 1h em TBS 1X (50 mM Tris, 150 mM NaCl, pH 8,0) contendo 5 % leite Molico desnatado (Nestlé). Após o bloqueio, foram feitas 3 lavagens de 5 min com TBS 1X e a membrana foi incubada com o anticorpo primário diluído 1:2500 em TBS 1X. Em seguida, a membrana foi lavada como descrito anteriormente e incubada por 90 min com o anticorpo

secundário anti-mouse-HRP do kit *Amersham ECL Western Blotting System* (GE), numa diluição 1:10000. A membrana foi lavada e a interação antígeno anticorpo foi revelada utilizando uma mistura de 500  $\mu\text{L}$  do reagente de detecção 1 e 500  $\mu\text{L}$  do reagente de detecção 2 presentes no mesmo kit.

### **3.6. Caracterização cinética da enzima recombinante CaneLEG2**

#### **3.6.1. Cálculo da constante de Michaelis Menten ( $K_m$ )**

A constante Michaelis-Menten ( $K_m$ ) permite analisar a afinidade da enzima pelo substrato utilizado. Para determinar o valor de  $K_m$ , 0,34  $\mu\text{M}$  da enzima foi pré-incubada por 5 minutos em 450  $\mu\text{L}$  de solução de reação contendo tampão citrato-fosfato pH 6,4 e DTT 2,5 mM. O DTT é utilizado como agente redutor do resíduo de cisteína presente no sítio ativo das cisteíno-peptidases. Em seguida foram adicionadas concentrações crescentes dos substratos Z-Val-Ala-Asn-MCA ou Z-Ala-Ala-Asn-MCA variando de 0,5 a 256  $\mu\text{M}$ , em 450  $\mu\text{L}$ . A velocidade de reação foi medida considerando a quantidade de MCA (7-amino-4-metil coumarina) livre, proveniente da hidrólise do substrato pela enzima. A molécula de MCA livre quando excitada em um comprimento de onda (  $\lambda$  ) de 360 nm emite fluorescência que pode ser detectada em 480 nm em espectrofluorímetro F-2500 (Hitachi). Desta forma, o aumento na fluorescência é proporcional à velocidade de reação. A inclinação das retas geradas no programa FL Solutions 2.0 a partir dos dados UAF vs tempo correspondem à velocidade de reação. Cada medida foi realizada em triplicata. O valor de  $K_m$  foi determinado a partir de um gráfico da velocidade de reação vs [substrato], utilizando o programa GraphPad 6.0.

#### **3.6.2. Ensaio de inibição enzimática**

Nos estudos de inibição da atividade da CaneLEG2 foram usados os inibidores: Ac-YVAD-CHO (Bachem) (inibidor de caspase-1); CaneCPI-3 (cistatina de cana-de-açúcar), CaneCPI-3DeIC (região N-terminal da CaneCPI-3) e a CaneCPI-3CT, extensão C-terminal da fitocistatina 3 de cana-de-açúcar caracterizada por Gianotti *et al.* (2006) que apresenta motivo inibitório de leguminaína. Nos ensaios de inibição enzimática, 0,34  $\mu\text{M}$  da CaneLEG2 foi pré-incubada por 5 min a 37°C em 450  $\mu\text{L}$  de solução de reação contendo tampão

citrato-fosfato pH 6,4 e DTT 2,5 mM. Após a incubação, 13,4  $\mu$ M do substrato Z-Val-Ala-Asn-MCA foi adicionado e foi realizada a leitura da fluorescência emitida por 150 segundos em espectrofluorímetro F-2500 (Hitachi). A atividade da enzima na ausência de inibidores, representada como 100% de atividade, foi utilizada como controle. Em seguida concentrações crescentes dos inibidores foram adicionadas a cada 100 segundos e foram feitas leituras das atividades residuais.

A constante de inibição ( $K_i$ ) foi calculada baseada na fórmula  $v_0/v_i = 1 + [I]_0/K_i$ , onde  $V_0$  é a velocidade de reação na ausência do inibidor,  $V_i$  a velocidade na presença do inibidor e  $[I]$  corresponde à concentração de inibidor utilizada. Após a obtenção dos valores de atividade da enzima na ausência e presença de inibidores, foi desenhado um gráfico  $(v_0/v_i) - 1$  vs  $[I]_0$  cuja inclinação da reta corresponde a  $1/K_i$  (BEYNON; BOND, 1990).

## 4. RESULTADOS E DISCUSSÃO

### 4.1. Análise *in silico*

A busca nas sequências do NCBI referentes ao banco de dados do SUCEST (*Sugarcane Expressed Sequence Tag*) utilizando como modelo a sequência da CaneLEG, resultou na identificação de 93 sequências, referentes a 87 clones diferentes. O número de sequências encontradas em cada biblioteca que compõe o SUCEST pode ser visualizado na Tabela 4.1.

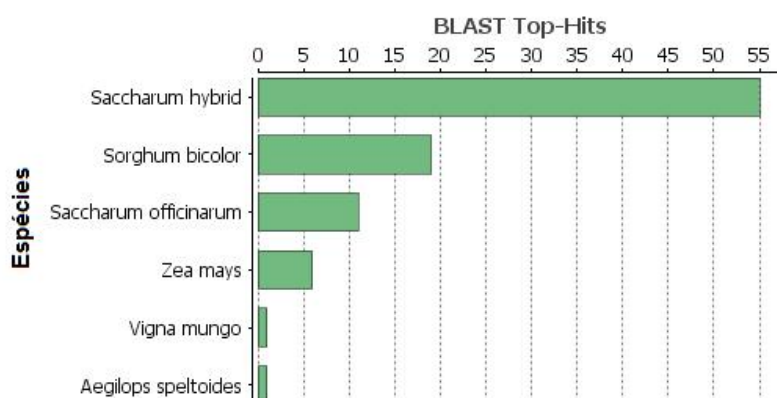
**Tabela 4.1. Descrição resumida das bibliotecas que compõem o SUCEST.** Na quarta coluna, encontra-se o número de sequências de legumaínas de cana-de-açúcar encontradas. Tabela modificada de VETTORE *et al.*, 2001.

Código da biblioteca	Nome da biblioteca	Descrição	N ° de Sequências
AD1	<i>G. diazotrophicans</i>	Tecidos de raiz, colmo e meristema apical infectados com <i>Gluconacetobacter diazotrophicans</i>	8
AM1, AM2	Meristema Apical	Meristema apical de plantas jovens	5
CL6	Calo	Pool de calos tratados por 12 h a 4°C e 37°C, no escuro ou na luz	1
FL1, FL3, FL4, FL5, FL8	Flores 1, 3, 4, 5 e 8	Flores em diferentes estágios de desenvolvimento	8
HR1	<i>H. rubrisubalbicans</i>	Tecidos de raiz, colmo e meristema apical de plantas cultivadas <i>in vitro</i> e infectados com <i>Herbaspirillum diazotrophicans</i>	5
LB1, LB2	Gema lateral 1 e 2	Gemas lateral de plantas maduras.	10
LR1, LR2	Cartucho foliar 1 e 2	Cartucho foliar de plantas imaturas	6
LV1	Folha 1	Folhas estioladas de plântulas crescendo <i>in vitro</i>	2
RT1, RT2, RT3	Raiz 1, 2 e 3	Raízes de 0,3 cm de plantas maduras e ápice de raiz	11
RZ1, RZ2, RZ3	Raiz ao ápice	Zonas da raiz ao ápice de plantas jovens	3
SB1	Casca do caule	Superfície do caule de plantas maduras	3
SD1, SD2	Sementes 1 e 2	Sementes em desenvolvimento	13
ST1, ST3	Haste 1 e 3	Primeiro e quarto entrenós de plantas imaturas	18

Representantes de legumaínas foram encontrados em todas as bibliotecas, o que corrobora com os resultados encontrados por Santos-Silva *et*

al. (2012), que investigaram a expressão tecidual da CaneLEG e observaram que o gene dessa enzima é expresso constitutivamente em todos os tecidos analisados (folhas, cartucho foliar, raízes, gema, meristema apical, entrenó e broto lateral) e nas quatro fases de desenvolvimento em que os tecidos foram coletados (estabelecimento, perfilhamento, crescimento dos colmos e maturação).

As sequências foram submetidas à análise no programa Blast2GO (CONESA; GÖTZ *et al.*, 2005) que serviu como uma ferramenta de anotação para confirmar a similaridade das sequências encontradas no SUCEST com legumaínas de outros bancos de dados. Apenas duas sequências não apresentaram similaridade com legumaína e foram eliminadas do restante das análises. Todas as demais mostraram alta similaridade com legumaínas de cana-de-açúcar, sorgo e milho (Figura 4.1).



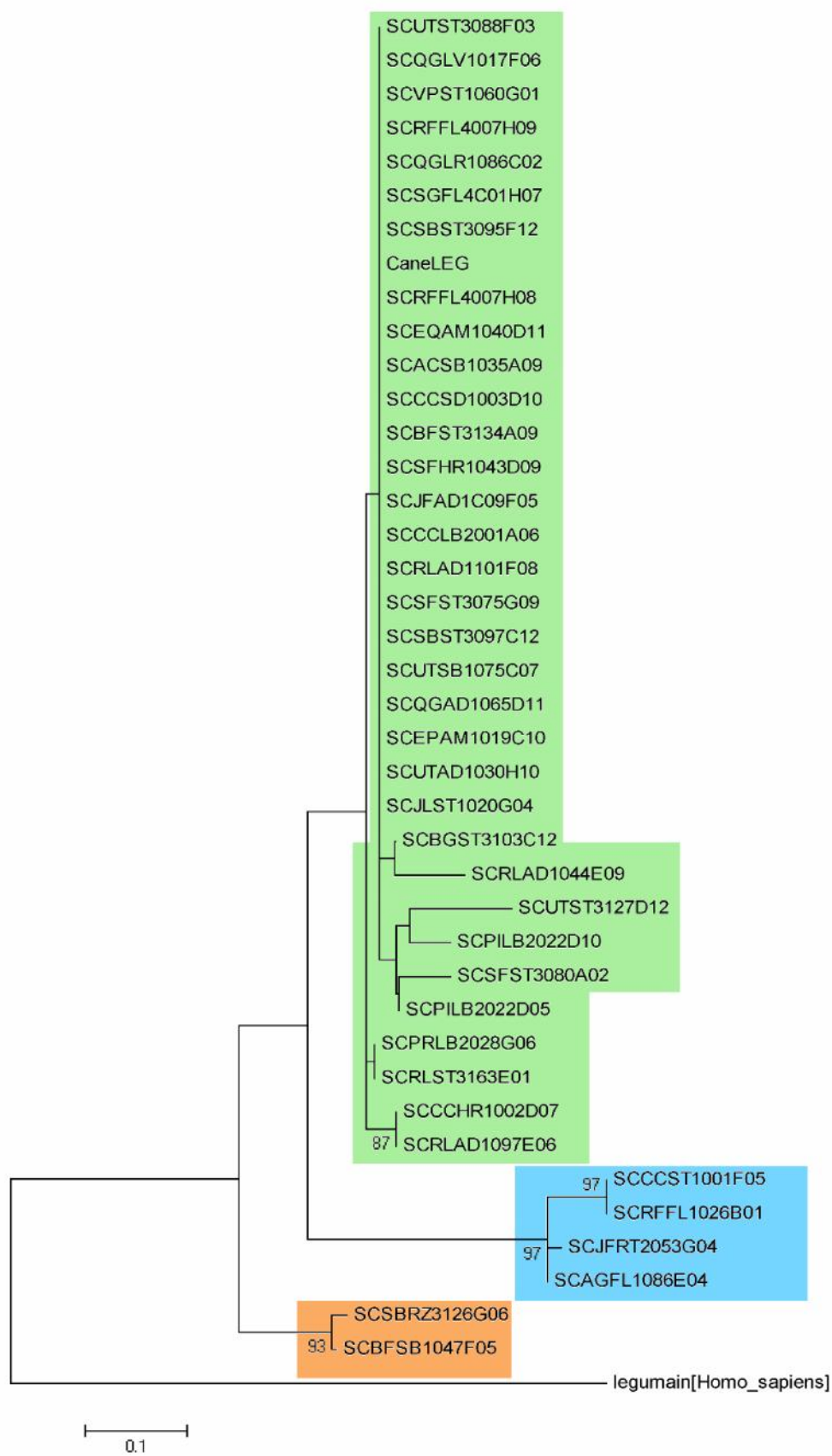
**Figura 4.1.** Gráfico obtido no programa Blast2GO mostrando a distribuição de espécies com maior similaridade encontradas na análise.

As sequências foram classificadas segundo a nomenclatura da IUBMB (*International Union of Biochemistry and Molecular Biology*) com o código E.C: 3.4.22 de cisteíno-endopeptidases. Os resultados do *Gene Ontology* associaram todas as sequências com o processo biológico de proteólise e função molecular de cisteíno-endopeptidase. A análise com *software* Pfam mostrou que todas pertencem à família de peptidases C13 e ao clan CD. Uma análise utilizando o *software* SignalP mostrou que as sequências possuem prováveis peptídeos sinais, característica típica das enzimas dessa família.

Correa *et al.* (2001) descreveram 22 *clusters* diferentes de legumaínas encontrados em cana-de-açúcar, porém uma análise mais detalhada desses

dados revelou que vários deles contêm sequências iguais, ou não apresentam similaridade com legumaínas dos bancos de dados. Portanto, para verificar se existem tipos diferentes de legumaínas na cana-de-açúcar, foi realizada uma análise filogenética com as sequências obtidas para classificá-las em grupos distintos, levando em consideração as diferenças entre elas. Primeiramente foi realizado um alinhamento, mas devido ao tamanho reduzido ou baixa qualidade de algumas sequências, apenas 59 foram usadas. Por meio desse alinhamento foi possível observar que 39 sequências correspondiam aos aminoácidos da região N-terminal da legumaína, enquanto 20 correspondiam à região C-terminal. Isso se deve à extensão da ORF completa da legumaína, que é superior à quantidade de nucleotídeos sequenciados para cada clone e depositados no banco de dados. Com a impossibilidade de realizar um alinhamento com todas elas, foi construída uma árvore filogenética para a região N-terminal, que possuía uma amostragem maior. A sequência da CaneLEG (GenBank ID: JQ083602) foi utilizada como representante de um dos grupos e a sequência de uma legumaína humana (NP\_005597.3) foi usada como controle do tipo *outgroup*. A análise da árvore filogenética resultante levou à classificação em três grupos (Figura 4.2).

O maior grupo foi designado grupo I, ao qual pertence à legumaína já caracterizada no próprio laboratório, CaneLEG. O clone SCJFRT2053G04, foi selecionado como representante do grupo II e foi comprado no Centro Brasileiro de Estocagem de Genes – BCC Center (Jaboticabal) e re-sequenciado.



**Figura 4.2.** Árvore filogenética construída com o programa MEGA 5 para clusterização das sequências de legumaínas encontradas no SUCEST. O grupo I está representado em verde, o II em azul e o III em laranja.



Com o re-sequenciamento do clone SCJFRT2053G04, foi possível obter uma sequência de cDNA contendo a ORF completa de uma legumaina de cana-de-açúcar diferente da CaneLEG já caracterizada, denominando assim a nova como CaneLEG2, contendo 1482 pb codificando para uma proteína de 493 aminoácidos (Figura 4.3).

UTR 5'

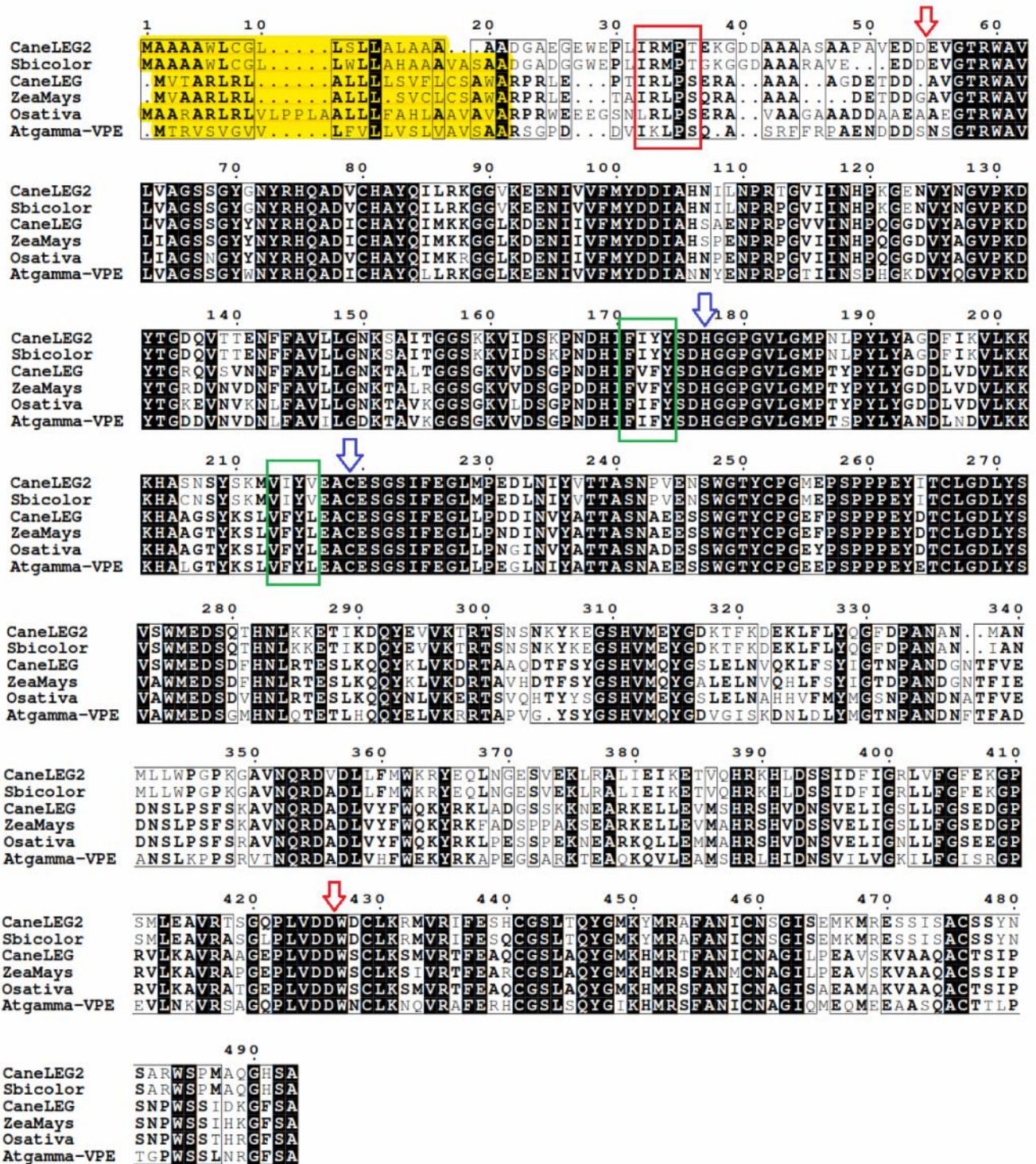
cacgcgtacgtaagccttgggatcctctaagcggccgcccgcgcggggccatgctgcggttgcgcegcac  
gacacgctcacttcgggtcgtatcttaaggcgcggccggaccagcctcgccgtcaccactcacc  
cactctctcgcacctaccgcctgctctctctgtctgacgcgcggccctcgcgatcggtg

1 ATGGCAGCTGCGGCG TGGCTGTGTGGACTC CTCTCGCTCCTGGCC CTCGCCGCGGC GCGGATGGGGCGGAG  
1 M A A A A W L C G L L S L L A L A A A A A D G A E  
76 GGTGAGTGGGAGCCG TTGATACGGATGCCG ACGGAGAAGGGGAT GACGCTGCTGCTGCT TCTGCTGCTCCGGCG  
26 G E W E P L I R M P T E K G D D A A A A S A A P A  
151 GTTGAGGACGATGAA GTGGGACGAGGTGG GCCGTGCTCGTCGCC GGCTCCTCCGGCTAC GGGAACTACCGGCAC  
51 V E D D E V G T R W A V L V A G S S G Y G N Y R H  
226 CAGGCCGATGTGTGC CATGCGTACCAGATA CTGCGGAAGGGAGGA GTGAAGGAGGAAAAC ATTGTGGTGTTCATG  
76 Q A D V C H A Y Q I L R K G G V K E E N I V V F M  
301 TATGATGATATTGCC CATAACATTTTGAAC CCGAGGACCGGGGTC ATCATTAAACCATCCT AAAGGCGAAAACGTC  
101 Y D D I A H N I L N P R T G V I I N H P K G E N V  
376 TATAATGGTGTTC AAGGATTACACAGGT GATCAGGTCCTACT GAAAACCTTCTTCGCT GTCCTCTTGGGCAAT  
126 Y N G V P K D Y T G D Q V T T E N F F A V L L G N  
451 AAAAGTGTCTACT GGAGGAGTAAGAAG GTGATAGACAGCAA CCTAATGACCACATC TTCATCTATTACTCA  
151 K S A I T G G S K K V I D S K P N D H I F I Y Y S  
526 GATCATGGGGTCTT GGAGTTCTTGGTATG CCAAACCTTGCCATAT CTGTATGCTGGCGAC TTCATCAAGGTTTTG  
176 D H G G P G V L G M P N L P Y L Y A G D F I K V L  
601 AAAAGAAGCATGCT TCCAATAGCTACTCA AAAATGGTTATATAT GTTGAAGCATGTGAA AGTGGCAGTATCTTC  
201 K K K H A S N S Y S K M V I Y V E A C E S G S I F  
676 GAGGATTAATGCCT GAAGTCTAAATAT TATGTCACACAGCC TCTAATCCAGTTGAA AATAGTTGGGGAACA  
226 E G L M P E D L N I Y V T T A S N P V E N S W G T  
751 TACTGCCCTGGGATG GAACCATCACCCCT CCTGAGTACATTACC TGTTTAGGTGATTTG TACAGTGTTCCTTGG  
251 Y C P G M E P S P P P E Y I T C L G D L Y S V S W  
826 ATGGAAGACAGTCAA ACCCAATCTAAAG AAGGAAACCATCAAG GATCAGTATGAAGTG GTGAAAACAAGAACC  
276 M E D S Q T H N L K K E T I K D Q Y E V V K T R T  
901 TCGAACTCGAATAAG TACAAAAGAGGTTCT CATGTTATGGAGTAT GGTGATAAGACCTTC AAGGATGAGAAGCTT  
301 S N S N K Y K E G S H V M E Y G D K T F K D E K L  
976 TTCCTTTATCAAGGT TTTGATCCTGCTAAT GCCAACATGGCAAAC ATGCTGCTTTGGCCA GGCCCAAAGGGTGCA  
326 F L Y Q G F D P A N A N M A N M L L W P G P K G A  
1051 GTCAATCAAAGAGAT GTTGATCTTCTCTTC ATGTGGAAGAGGTAT GAACAGTTAAATGGG GAGTCTGTGGAGAAG  
351 V N Q R D V D L L F M W K R Y E Q L N G E S V E K  
1126 CTGAGGGCTCTCATA GAGATTAAAGAACT GTTCAACATAGGAAG CATCTTGACAGCAGT ATCGATTTCAATTGGG  
376 L R A L I E I K E T V Q H R K H L D S S I D F I G  
1201 AGGCTTGTCTTCGGA TTTGAGAAAGGGCCT TCGATGCTCGAGGCT GTTAGAACCTCGGGT CAACCATGGTTCGAT  
401 R L V F G F E K G P S M L E A V R T S G Q P L V D  
1276 GACTGGGATTGTTT AAGAGGATGGTGCAG ATTTTCGAATCCCAT TGTGGTTCACTTACT CAGTATGGCATGAAG  
426 D W D C L K R M V R I F E S H C G S L T Q Y G M K  
1351 TACATGAGAGCGTTT GCAAACATTTGCAAC AGTGGTATATCTGAG ATGAAAATGAGGGAG TCAAGCATCAGCGCT  
451 Y M R A F A N I C N S G I S E M K M R E S S I S A  
1426 TGCAGCAGCTACAAC TCGGCGAGGTGGAGC CCGATGGCTCAGGGA CACAGCGCTTGA  
476 C S S Y N S A R W S P M A Q G H S A \*

tgatgataagatcaatagcaaaagcccccgaggtctgctgccccctgtgctgtgtgagctcgta  
tcatgtggggactgtgaaccatcgtgtacttatgttctgtacctatgtaaagtaatgccaagtctga  
tcttcatccaactaccgaaatgattaagatcagatgtgaatggtgttgcgatggatgatcgcaagtt  
gctataatcagtggaacaaaaatgataacccccaaaaaaa UTR 3'

**Figura 4.3.** Sequência de nucleotídeos e proteína predita da CaneLEG2. A região sublinhada representa o peptídeo sinal com a seta indicando o sítio de clivagem. O círculo indica o provável sítio de glicosilação. As letras minúsculas correspondem às regiões não traduzidas (UTR).

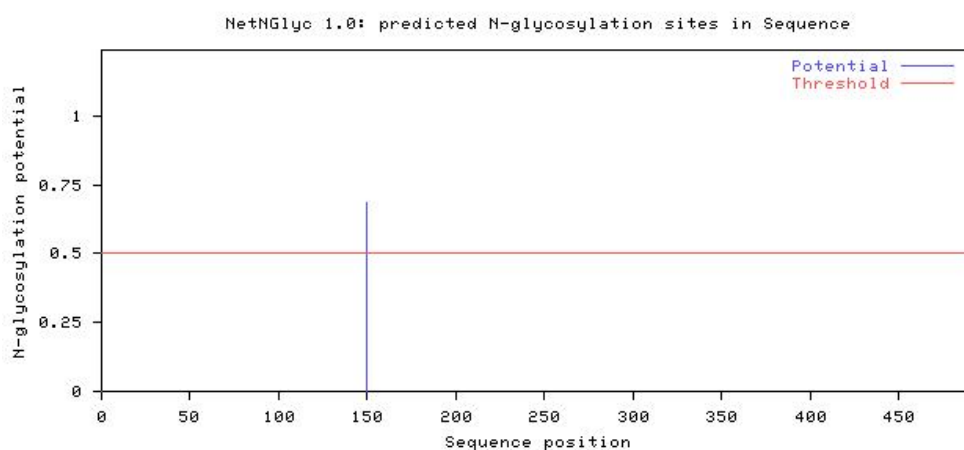
A figura 4.4 mostra o alinhamento da sequência obtida deste clone para CaneLEG2 com CaneLEG e legumainas de outros organismos.



**Figura 4.4.** Alinhamento entre CaneLEG2 do SUCEST e outras legumainas pelo programa Multialin. O alinhamento de aminoácidos foi realizado com outras VPEs dos organismos *Saccharum sp.* (JQ083602.1), *Arabidopsis thaliana* (BAA18924.1), *Sorghum bicolor* (XP\_002448237.1), *Zea Mays* (CAB64544.1) e *Oryza sativa* (BAA84650.1). Os blocos pretos representam as sequências idênticas de aminoácidos e aminoácidos similares estão em negrito. Os peptídeos sinais estão em amarelo. A caixa vermelha indica o motivo descrito por Jackson *et al.* (2007). As setas vermelhas mostram os prováveis sítios de clivagem dos pró-peptídeos N e C-terminal. Os blocos verdes indicam os quatro aminoácidos que precedem os resíduos catalíticos, que estão indicados pelas setas azuis.

A CaneLEG2 possui alta identidade com VPEs de outras plantas, com regiões conservadas, como os resíduos catalíticos, Hys e Cys, além dos prováveis sítios de clivagem dos pró-peptídeos N e C-terminal. Entretanto, existem algumas diferenças que a separam da CaneLEG, como a composição dos quatro aminoácidos hidrofóbicos que precedem os resíduos catalíticos, entre outras diferenças ao longo da sequência, em regiões menos conservadas. Além disso, o motivo IRMPT da CaneLEG2 difere do motivo conservado em leguminais de plantas descrito por Jackson *et al.* (2007) (I/L)(R/K)(L)(P)(S), o qual está relacionado ao direcionamento da legumina para vacúolos líticos.

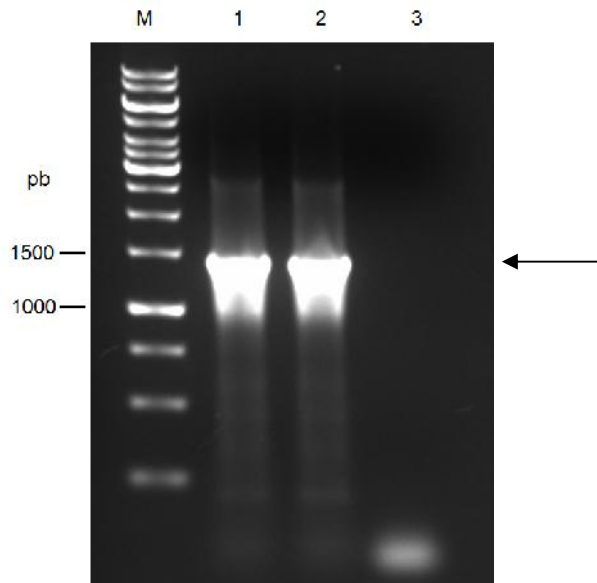
A análise da sequência de aminoácidos da CaneLEG2 no programa Signal P (BENDTSEN *et al.*, 2004) indicou que os 19 aminoácidos iniciais correspondem a um possível peptídeo sinal, provavelmente envolvido no direcionamento da proteína para o retículo endoplasmático, onde ela entra na via de secreção e é encaminhada ao vacúolo, de acordo com análise no programa de predição de direcionamento celular PSORT (HORTON *et al.*, 2007). O provável sítio de glicosilação predito com o programa NetNGlyc 1.0 pode ser observado na figura 4.5 e encontra-se no resíduo de asparagina na posição 150, na sequência da CaneLEG2.



**Figura 4.5.** Provável sítio de glicosilação da CaneLEG2. O gráfico gerado com o software NetNGlyc 1.0 mostra um possível sítio de N-glicosilação na asparagina 150 da cadeia de aminoácidos da CaneLEG2.

#### 4.2. Clonagem da CaneLEG2 no vetor de expressão pGAPZ A

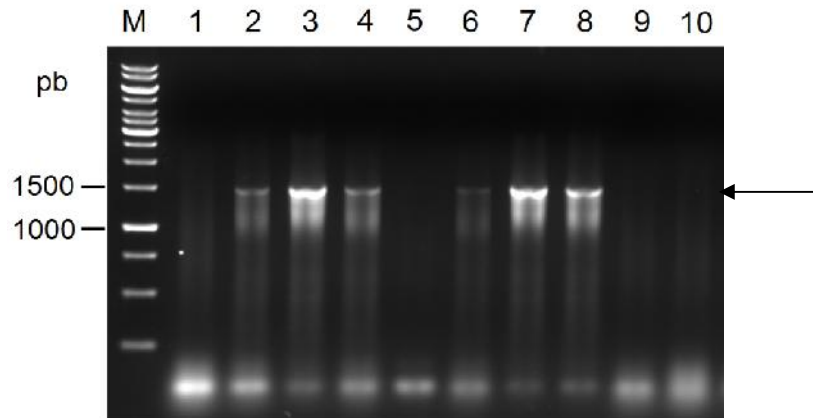
A ORF correspondente à CaneLEG2, sem o peptídeo sinal e sem códon de terminação, foi amplificada por PCR utilizando os oligonucleotídeos com os sítios para as enzimas de restrição. A figura 4.6 mostra a análise em gel de agarose 1 %, onde se observa a presença do fragmento no tamanho esperado de 1439 pb.



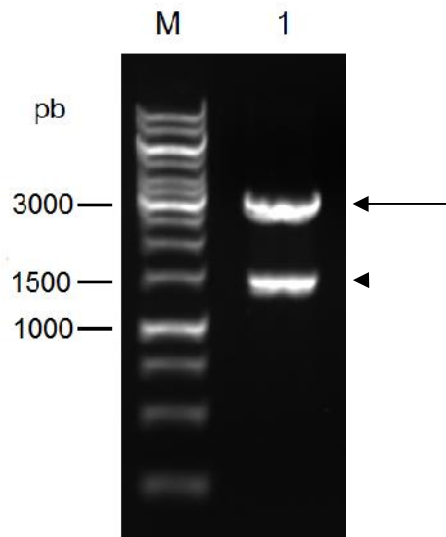
**Figura 4.6.** Amplificação da ORF da legumaina de cana-de-açúcar (CaneLEG2) a partir do clone SCJFRT2053G04, utilizando *primers* específicos. Gel de agarose 1% mostrando em M: marcador de massa molecular 1kb DNA Ladder (Invitrogen), 1 e 2: produto amplificado em duplicata, indicado pela seta, 3: controle negativo.

O amplicon foi digerido e clonado no vetor de expressão pGAPZ A, que permite a inserção da cauda de histidinas na região C-terminal da proteína recombinante. Para confirmação da inserção correta do fragmento, além da PCR das colônias (Figura 4.7), foi realizada uma análise de restrição com as enzimas utilizadas na clonagem (*Sall* e *EcoRI*). Após a digestão, a análise em gel de agarose mostrou a presença de duas bandas correspondentes ao vetor (3,1 kb) e ao inserto (1422 pb), como mostrado na figura 4.8.





**Figura 4.7.** PCR de colônia para identificação de transformantes de *E. coli* DH5 com o plasmídeo recombinante contendo a ORF da legumaína, mostrando em M: marcador de massa molecular 1kb DNA Ladder (Invitrogen), 1: controle negativo, 2 a 10: clones transformados com o plasmídeo recombinante. A seta indica a banda amplificada de aproximadamente 1500 pb.



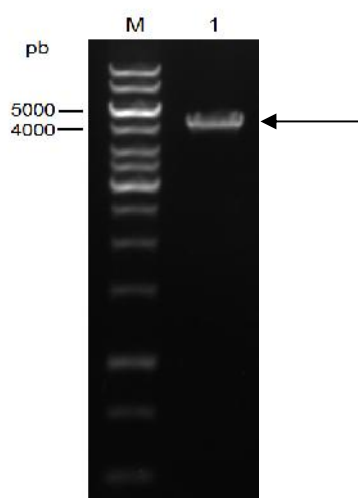
**Figura 4.8.** Análise de restrição do plasmídeo recombinante contendo a ORF da legumaína. Gel de agarose 1% mostrando em M: marcador de massa molecular 1kb DNA Ladder (Invitrogen), 1: plasmídeo recombinante clivado com as enzimas de restrição. A seta indica a banda de 3 kb correspondente ao vetor pGAPZ A e a cabeça de seta indica o inserto clonado de 1422 pb.

A construção contém a ORF que codifica a proteína na sua forma inativa, contendo os pró-peptídeos N e C-terminal, além da sequência sinal de secreção fator presente no vetor, proveniente de *Saccharomyces cerevisiae*, que permite a secreção da proteína no meio de cultivo. A levedura *P. pastoris* secreta uma baixa quantidade de proteínas endógenas, fazendo com que esse direcionamento para o meio de cultivo sirva como um passo inicial para a purificação. A levedura *P. pastoris* é um sistema eficiente para expressão heteróloga de proteínas que permite, quando comparado com expressão em

bactérias, um melhor enovelamento e processamento de proteínas eucarióticas e ainda torna possível a realização de glicosilação em proteínas secretadas, que não é obtida em sistemas de expressão procarióticos. Existe ainda a possibilidade de integração dos vetores de expressão ao genoma da levedura, aumentando a estabilidade das linhagens utilizadas. Para produção da proteína recombinante foi utilizada a linhagem de *P. pastoris* KM71H (his4 arg4 aox1 ::ScARG4).

A enzima gliceraldeído-3-fosfato desidrogenase (GAPDH) é constitutivamente expressa em altos níveis em muitos organismos, incluindo a *Pichia pastoris*. O promotor do gene (*GAP*) que codifica a GAPDH tem se mostrado eficiente na produção de altos níveis de proteínas recombinantes em *P. pastoris*, dependendo da fonte de carbono utilizada (WATERHAM *et al.*, 1997). O nível de expressão observado com o promotor *GAP* ( $P_{GAP}$ ) pode ser ligeiramente maior do que o obtido com o promotor AOX1 e dispensa o uso de metanol, que pode ser problemático em alguns casos (CREGG *et al.*, 2000). O promotor *GAP* também é conveniente, pois a cultura não precisa ser trocada de uma fonte de carbono para outra para induzir a expressão de um gene exógeno.

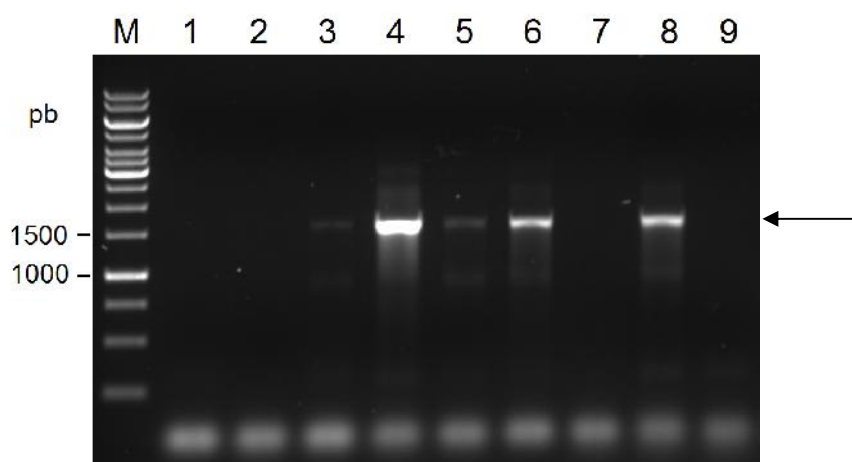
O plasmídeo pGAPZ A contendo a ORF da CaneLEG2 foi sequenciado e após confirmação da sequência o plasmídeo foi linearizado (Figura 4.9) e utilizado para transformar células eletrocompetentes de *P. pastoris* KM71H.



**Figura 4.9.** Plasmídeo recombinante pGAPZ A-CaneLEG2 linearizado. Gel de agarose 1 % mostrando em M: marcador de massa molecular 1 kb DNA Ladder (Invitrogen), 1: Precipitação do vetor recombinante linearizado, indicado pela seta.

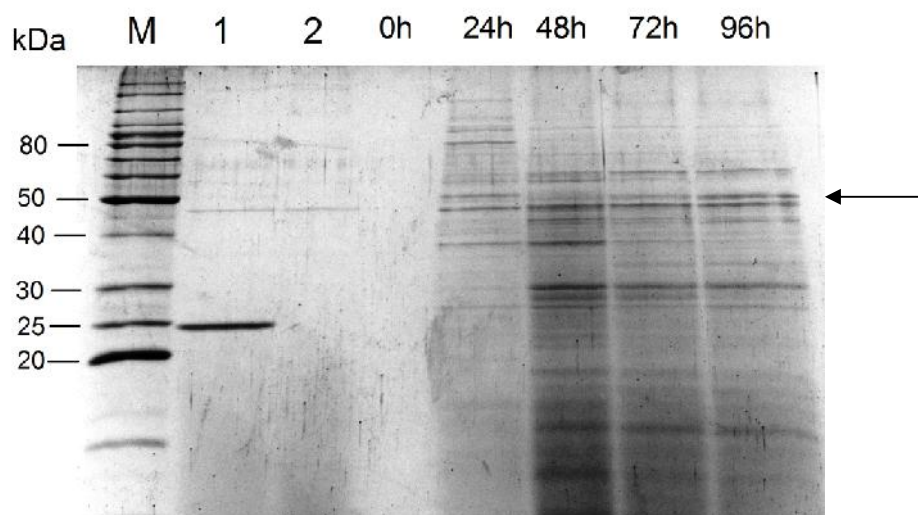
### 4.3. Transformação de células competentes de *P. pastoris* e expressão da CaneLEG2 recombinante

A confirmação da inserção do gene de interesse (CaneLEG2) no genoma da levedura foi feita por PCR das colônias dos clones que cresceram nas placas contendo meio seletivo. Dos 8 clones analisados, 5 apresentaram o produto de amplificação no tamanho esperado correspondente à ORF da CaneLEG2, de aproximadamente 1500 pb (Figura 4.10). A integração do vetor de expressão no genoma da levedura aumenta a estabilidade do transformante que, após seleção e confirmação do clone recombinante, não precisa ser mantido em meio seletivo.



**Figura 4.10.** PCR das colônias de KM71H transformadas com plasmídeo pPGAPZ A-CaneLEG2. Gel de agarose 1% mostrando em M: marcador de massa molecular 1kb DNA Ladder (Invitrogen), 1: controle negativo, 2 a 9: clones testados para a presença do plasmídeo recombinante. A seta indica a banda amplificada de aproximadamente 1500 pb.

Apesar do resultado positivo obtido na PCR das colônias 3, 4, 5, 6 e 8, ainda é necessário analisar a expressão da proteína recombinante nos clones positivos para confirmar se o gene foi inserido na região correta do genoma, portanto esses clones foram utilizados para o ensaio de expressão da proteína recombinante em pequena escala por 96 h. Foi observada a expressão da proteína recombinante em um dos clones (Figura 4.11), que foi selecionado e isolado para os ensaios posteriores.



**Figura 4.11.** Expressão da CaneLEG2 recombinante em *P. pastoris*. SDS-PAGE 12 % mostrando em M: marcador de massa molecular BenchMark (Invitrogen), 1: controle positivo (Endoglucanase 3 de *Trichoderma harzianum* clonada em pGAPZ A), 2: controle negativo (KM71H não transformada), as demais colunas mostram amostras do meio de indução do clone transformado com pGAPZ A-CaneLEG2. O tempo de indução é indicado na figura. A seta indica a provável banda da CaneLEG2 de aproximadamente 54 kDa contendo os pró-peptídeos N e C-terminal.

Foi possível observar a expressão da CaneLEG2 a partir de 24 h de indução, entretanto a proteína apresentava bandas de tamanhos variados, diferentes do tamanho esperado de 54 kDa. Isso provavelmente ocorreu devido ao seu processamento auto-catalítico ou glicosilação. Com o acúmulo de legumainas ativas no meio de cultura, ela se autoprocessa realizando a clivagem dos pró-peptídeos N-terminal e C-terminal, gerando os fragmentos menores que foram observados em SDS-PAGE. Bandas com massas moleculares menores também podem ser sinal de degradação. Já os fragmentos que apresentaram massa molecular maior do que a esperada possivelmente são resultado de glicosilação pela levedura, uma vez que a análise da sequência de aminoácidos da CaneLEG2 indicou a presença de um possível sítio de glicosilação no resíduo de asparagina 150.

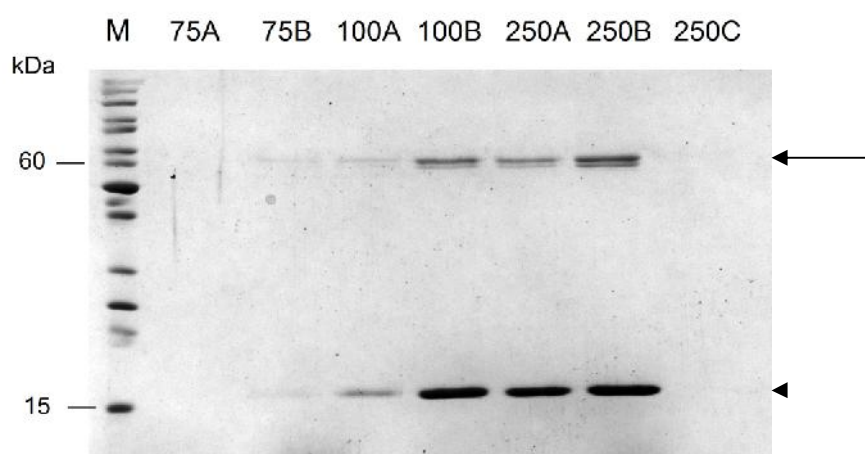
#### 4.4. Indução, expressão e purificação da CaneLEG2 recombinante

Foi realizada uma indução em 100 mL a fim de tentar obter a enzima purificada através de cromatografia de afinidade em coluna contendo resina de níquel. Porém a CaneLEG2 recombinante foi recuperada nas frações do eluato e da primeira lavagem da resina com tampão, indicando que houve perda do



pró-peptídeo C-terminal e, conseqüentemente, da cauda de histidinas essencial para a ligação da proteína à coluna de níquel. Para tentar evitar o processamento prematuro da enzima, foi feita uma nova indução com o pH alterado de 6,0 para 7,0, pois pesquisas demonstraram que legumainas de plantas possuem atividade ótima em pH ácido e que em condições neutras ou básicas sua atividade é baixa ou quase nula (ROTARI *et al.*, 2001; KUROYNAGI *et al.*, 2002; SANTOS-SILVA *et al.*, 2012). Entretanto, o processamento ainda foi observado e a purificação não foi possível (dados não mostrados).

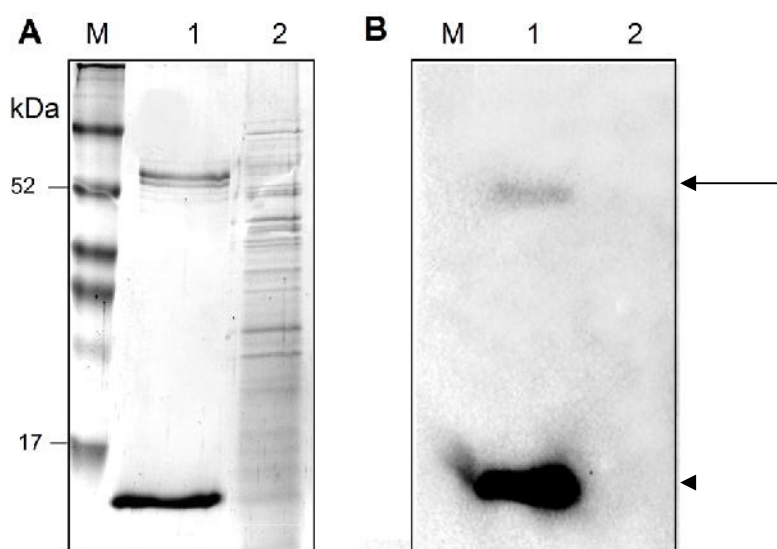
Para evitar que ocorresse um processamento precoce da CaneLEG2, ao inóculo foi adicionado 1 mg do inibidor CaneCPI-3CT. Esse inibidor consiste da região C-terminal de uma fitocistatina de cana-de-açúcar (CaneCPI-3) caracterizada por Gianotti *et al.* (2006), que apresenta motivo inibitório de legumaina descrito por Martinez *et al.* (2007). Dessa forma, o processamento da CaneLEG2 durante a indução foi impedido e a enzima permaneceu inativa, contendo os pró-peptídeos e, portanto, a cauda de histidinas na porção C-terminal. Isso tornou possível a purificação em coluna com resina de níquel, onde foram obtidas frações contendo a enzima em sua forma nativa (aproximadamente 60 kDa) e o inibidor (aproximadamente 15 kDa) (Figura 4.12). A massa molecular da CaneLEG2 apresenta-se um pouco maior no gel do que o tamanho esperado de 54 kDa, possivelmente devido a glicosilação.



**Figura 4.12.** Expressão heteróloga e purificação da CaneLEG2 junto com o inibidor CaneCPI-3CT. Análise em SDS-PAGE 12 % mostrando em M: marcador de massa molecular PageRuler (Fermentas) e nas demais colunas as frações com concentrações crescentes de imidazol, conforme indicado na figura. A seta indica a banda correspondente à CaneLEG2 e a cabeça de seta indica a banda do inibidor CaneCPI-3CT.

#### 4.5. Ensaio de imunodeteccção

Análise de imunodeteccção *Western Blotting* foi realizada para confirmar a enzima CaneLEG2, utilizando o anticorpo Anti-His em uma amostra da enzima induzida sem o inibidor e em uma amostra da enzima purificada junto com o inibidor CaneCPI-3CT (Figura 4.13). A CaneLEG2 foi detectada apenas na amostra purificada junto com a cistatina, enquanto na outra não houve sinal de reconhecimento do anticorpo, mostrando que houve inibição da atividade enzimática e seu processamento não ocorreu quando induzida com CaneCPI-3CT.

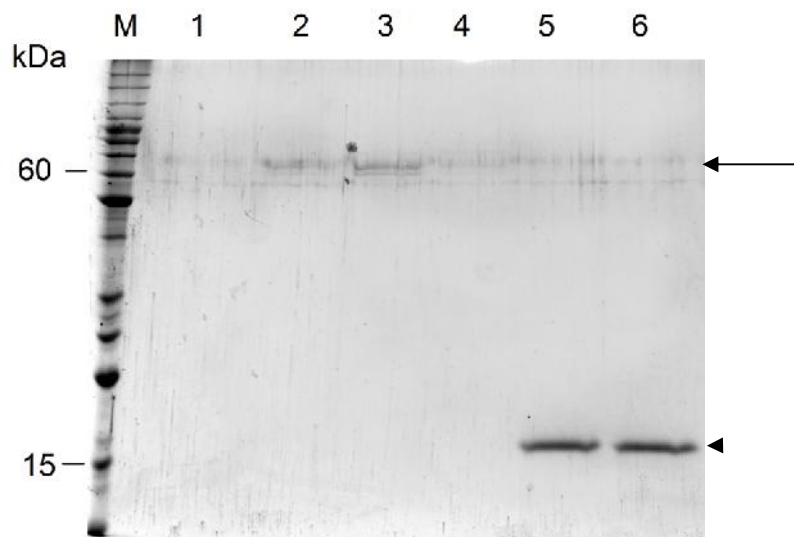


**Figura 4.13.** Detecção da CaneLEG2 recombinante produzida em *P. pastoris*. (A) SDS-PAGE 12 % corado com azul de comassie, mostrando em M: Marcador de massa molecular High-Range Rainbow (GE), 1: CaneLEG2 (~60 kDa) e CaneCPI-3CT (15 kDa) purificados, 2: CaneLEG2 na ausência de CaneCPI-3CT. (B) Membrana de PVDF réplica do gel mostrado em (A) utilizada para detecção da CaneLEG2 com anticorpo Anti-His. A seta indica a banda correspondente à CaneLEG2 e a cabeça de seta indica a banda do inibidor CaneCPI-3CT.

#### 4.6. Separação do complexo enzima-inibidor e ativação da CaneLEG2

As frações contendo a enzima CaneLEG2 e a cistatina CaneCPI-3CT foram dialisadas em tampão de lise pH 8,0, 1 M NaCl para que o complexo enzima-inibidor fosse dissociado devido ao aumento da força iônica. Uma nova etapa de purificação foi realizada utilizando uma coluna de gel filtração, Superdex™ 75 10/300 GL (GE Healthcare) possibilitando a separação

decorrente da diferença de tamanho entre a CaneLEG2 e a CaneCPI-3CT e então foram coletadas frações contendo a proteína recombinante pura, livre do inibidor e na sua forma inativa (Figura 4.14). O rendimento total da proteína recombinante purificada foi de aproximadamente 20 mg/L de meio de cultura YPD tamponado.



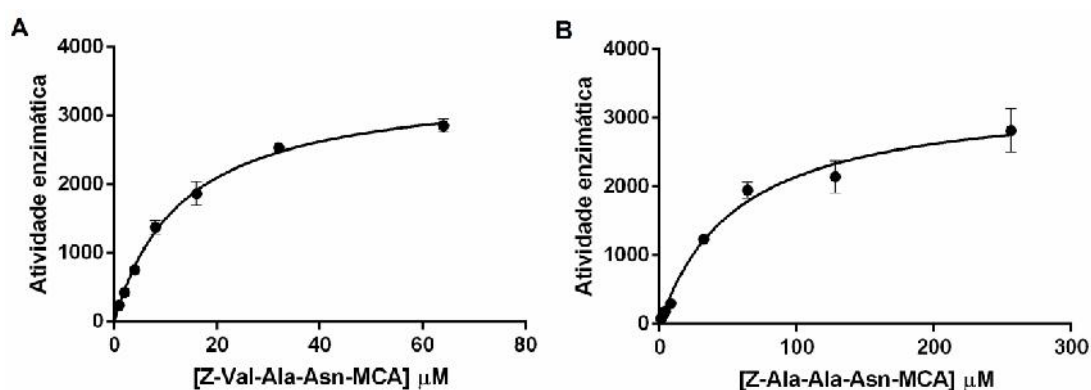
**Figura 4.14.** Purificação e separação da CaneLEG2 e o inibidor CaneCPI-3CT em coluna de gel filtração, Superdex™ 75 10/300 GL (GE Healthcare) utilizando o sistema de purificação de proteínas Akta FPLC (Amersham Pharmacia Biotech). Análise em SDS-PAGE 12 % mostrando em M: marcador de massa molecular BenchMark (Invitrogen), 1 a 4: frações contendo a CaneLEG2 purificada, 5 e 6: frações contendo o inibidor. A seta indica a banda correspondente à CaneLEG2 e a cabeça de seta indica a banda do inibidor CaneCPI-3CT.

A ativação da enzima foi realizada dialisando as frações purificadas contra tampão citrato-fosfato (138,5 mM  $\text{Na}_2\text{HPO}_4$ , 30,7 mM ácido cítrico, pH 6,4). Após a diálise e quantificação, foram realizados os ensaios de atividade.

#### **4.7. Determinação da constante de Michaelis Menten da enzima CaneLEG2**

Ensaio de atividade com a CaneLEG2 recombinante mostraram que ela é capaz de hidrolisar tanto o substrato Z-Ala-Ala-Asn-MCA como o Z-Val-Ala-Asn-MCA, portanto as constantes de Michaelis Menten foram determinadas para a enzima utilizando concentrações crescentes desses substratos (Figura 4.15). A constante de Michaelis pode ser definida como a afinidade da enzima pelo substrato e é numericamente igual à concentração de substrato

necessária para que a enzima atinja metade da velocidade máxima de reação ( $V_{m\acute{a}x}$ ). A  $V_{m\acute{a}x}$  é teoricamente alcançada quando a enzima foi saturada com uma concentração infinita de substrato. Estudos realizados por Santos-Silva *et al.* (2012) mostraram que a primeira legumaina de cana-de-açúcar descrita, CaneLEG, apresentou eficiências catalíticas para hidrólise dos substratos Z-Ala-Ala-Asn-MCA e Z-Val-Ala-Asn-MCA de  $33\text{mM}^{-1}\text{s}^{-1}$  e  $65\text{mM}^{-1}\text{s}^{-1}$ , respectivamente, indicando que a CaneLEG tem preferência pelo substrato contendo resíduo de valina na posição P3. O valor de  $K_m$  calculado para a CaneLEG2 foi de  $13,4\ \mu\text{M}$  para Z-Val-Ala-Asn-MCA e  $56,73\ \mu\text{M}$  para Z-Ala-Ala-Asn-MCA, mostrando que a CaneLEG2 também possui maior afinidade pelo primeiro substrato. O valor de  $K_m$  obtido para Z-Ala-Ala-Asn-MCA é semelhante a outras legumainas de porco (*Sus scropha*) e feijão comum (*Phaseolus vulgaris*), que apresentaram valores de  $K_m$  de  $50$  e  $56\ \mu\text{M}$ , respectivamente, para o mesmo substrato, o qual é comumente usado na caracterização de legumainas (ROTARI *et al.* 2001). Em comparação com o  $K_m$  da CaneLEG de  $11,5\ \mu\text{M}$ , descrito por Santos-Silva *et al.* (2012), a CaneLEG2 mostrou ter uma afinidade menor com esse substrato.



**Figura 4.15.** Caracterização cinética da atividade enzimática da CaneLEG2 recombinante. A curva de Michaelis-Menten foi obtida pela velocidade da reação, indicada no eixo das ordenadas e a concentração crescente do substratos Z-V-A-N-MCA (A) e Z-A-A-N-MCA (B), encontrados no eixo das abscissas.

A enzima recombinante foi considerada estável, pois manteve sua atividade catalítica constante durante todos os ensaios.

#### 4.8. Ensaios de inibição da atividade da CaneLEG2 *in vitro*

Inibidores com mecanismo de ação ou proteínas-alvo conhecidos são comumente utilizados como uma ferramenta importante para auxiliar na caracterização de enzimas quanto a sua provável função e seu mecanismo de ação. A constante de inibição ( $K_i$ ), também conhecida como constante de dissociação do complexo enzima-inibidor, expressa a afinidade do inibidor pela enzima (CHENG; PRUSOFF, 1973). Um inibidor é considerado potente se sua afinidade pela enzima alvo for tão alta que o seu valor de  $K_i$  é menor ou igual à concentração total de enzima utilizada no ensaio de inibição (COPELAND, 2005).

Uma das características que separam as legumainas das cisteíno-peptidases do Clan C1 é o fato que as legumainas não sofrem inibição pelo E-64, um inibidor sintético típico de cisteíno-peptidases do tipo papaína. A CaneLEG2, quando testada com esse inibidor, não sofreu alteração na sua velocidade de reação (Tabela 4.2), corroborando com outros trabalhos que mostraram o mesmo comportamento em outras legumainas (KEMBHAVI *et al.*, 1993; ALIM *et al.*, 2007; SANTOS-SILVA *et al.*, 2012).

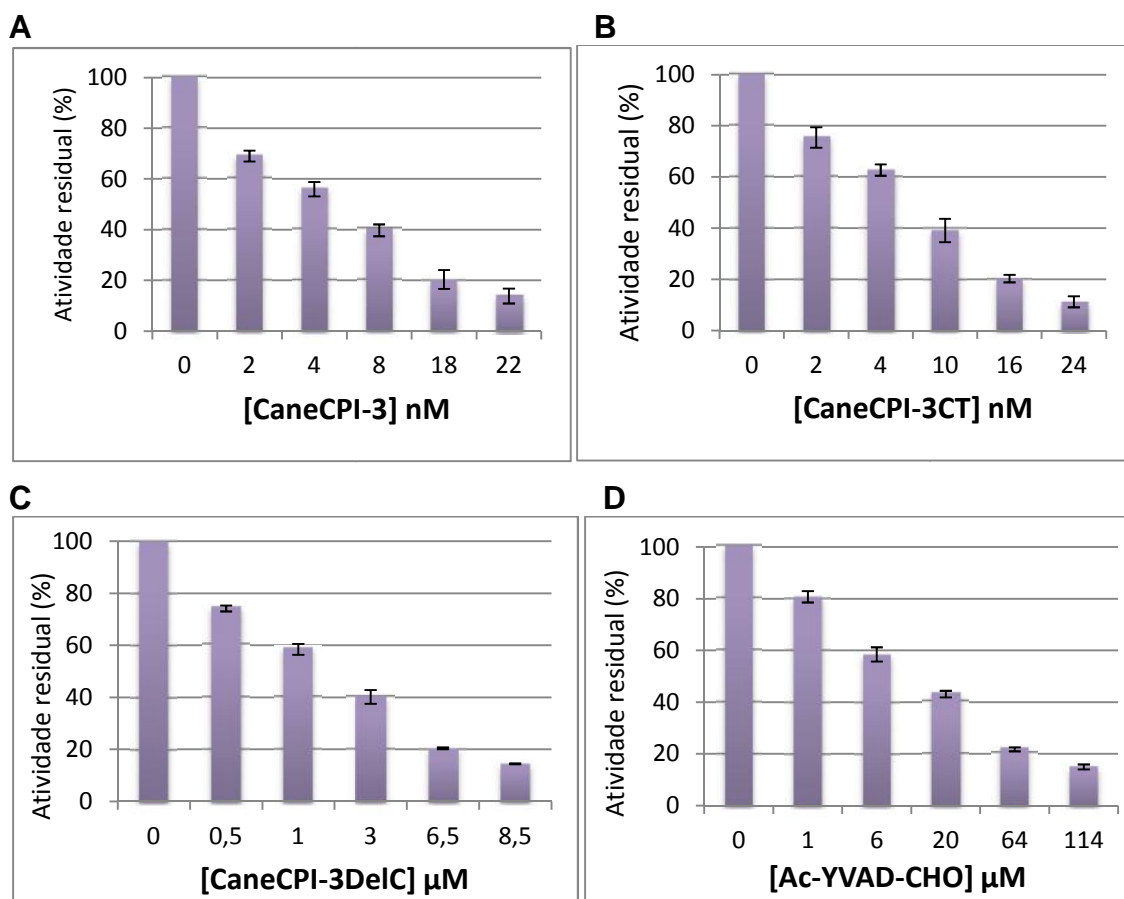
**Tabela 4.2.** Valores das constantes de inibição ( $K_i$ ) de diferentes inibidores contra a CaneLEG2.

Inibidor	$K_i$ (nM)
CaneCPI-3CT	4,22
CaneCPI-3	4,62
CaneCPI-3DelC	$1,87 \times 10^3$
Ac-YVAD-CHO	$20,83 \times 10^3$
E-64	n.i.*

\*não inibiu

Outro inibidor testado contra a CaneLEG2 foi a CaneCPI-3, uma fitocistatina de cana-de-açúcar, descrita por Gianotti *et al.* (2006). Santos-Silva *et al.* (2012) realizaram ensaios de inibição com essa fitocistatina e seus resultados mostram que ela é um forte inibidor da CaneLEG, com um  $K_i$  de 0,28 nM. Quando testada contra CaneLEG2, a CaneCPI-3 também apresentou forte inibição, com uma constante de inibição na ordem de nanomolar ( $K_i = 4,62$  nM) (Tabela 4.2, Figura 4.16 A). Em plantas, Martinez *et al.* (2007) estudaram uma cistatina de cevada de 23 kDa e encontraram a presença de

uma região C-terminal estendida capaz de inibir legumaina humana, bem como a atividade de legumainas em extratos proteicos de cevada. Essa região C-terminal é similar à cistatina M/E humana, que é capaz de inibir legumaina em um sítio reativo diferente do sítio de inibição de enzimas do tipo papaína (CHENG *et al.*, 2006). A CaneCPI-3 é muito similar à cistatina-4 de cevada (HvCPI-4) e já foi descrita sua atividade inibitória contra enzimas do tipo papaína (GIANOTTI *et al.*, 2006; MARTINEZ *et al.*, 2009). Santos-Silva *et al.* (2012) demonstraram que a inibição da CaneLEG pela CaneCPI-3 deve-se provavelmente à asparagina do motivo SNSL presente na extensão C-terminal, pois quando o resíduo de asparagina foi substituído em mutantes da CaneCPI-3 houve uma diminuição da atividade inibitória. Baseado nesses dados, foram realizados ensaios de inibição da CaneLEG2 com a região C-terminal estendida (CaneCPI-3CT) e com a região N-terminal (CaneCPI-3DelC). Já havia sido observado durante a etapa de purificação que a região C-terminal estendida, usada na indução da expressão, inibiu a atividade enzimática da CaneLEG2. A constante de inibição encontrada para a CaneCPI-3CT foi também na ordem de nanomolar ( $K_i = 4,22 \text{ nM}$ ) (Tabela 4.2, Figura 4.16 B). Como era esperado, a porção N-terminal da CaneCPI-3, por não possuir o motivo inibitório SNSL teve uma inibição bem menor, na escala de micromolar ( $K_i = 1,87 \text{ }\mu\text{M}$ ) (Tabela 4.2, Figura 4.16 C).



**Figura 4.16.** Inibição da CaneLEG2 por CaneCPI-3 (A) , CaneCPI-3CT (B), CaneCPI-3DelC (C) e Ac-YVAD-CHO (D). A atividade residual da enzima foi calculada a partir da inclinação das retas geradas com o programa FL solutions 2.0 e expressa como porcentagem de atividade em relação ao controle (ausência de inibidor).

Acredita-se que as cistatinas com massa molecular de 23 kDa encontradas em algas verdes e táxons de vegetais superiores foram resultado de um evento de duplicação gênica no início da evolução do grupo *Viridiplantae*, muito antes da irradiação adaptativa das plantas terrestres (MARGIS-PINHEIRO *et al.*, 2008; MARTINEZ *et al.*, 2005). O domínio C-terminal dessas proteínas teria passado por um intenso remodelamento estrutural e funcional, gradualmente levando à perda de atividade contra peptidases do tipo papaína e adquirindo novas funções, incluindo atividade inibitória contra cisteíno-peptidases do tipo legumaína (família C13) (MARTINEZ *et al.*, 2007) e um possível papel regulatório na função inibitória da porção N-terminal da cistatina (GIRARD *et al.*, 2007; WANG *et al.*, 2008).

Existem evidências de que haja uma regulação endógena por inibidores de peptidases e que um desequilíbrio desse controle de proteólise pode afetar diversos aspectos no desenvolvimento da planta (SOLOMON *et al.*, 1999; SIN;

YEUNG; CHYE, 2006; RIVARD *et al.*, 2007;). Isso indica que a CaneCPI-3 é um inibidor endógeno da CaneLEG e da CaneLEG2, provavelmente agindo na regulação da atividade enzimática na planta. Estudos demonstram que VPEs agem na mobilização de proteínas durante a germinação de sementes e crescimento da planta através da digestão direta dessas proteínas ou via processamento de outras endopeptidases (SCHLERETH *et al.*, 2000; OKAMOTO; MINAMIKAWA, 1999). Esse resultado é consistente com o fato de que a germinação na cana-de-açúcar é um estágio ativo em que ocorre a mobilização de nutrientes do broto em desenvolvimento (DILLEWIJN, 1952). Ademais, a germinação do broto lateral pode ser semelhante à germinação da semente, visto que proteínas de armazenamento previamente depositadas nos vacúolos são mobilizadas durante a germinação e usadas no crescimento da plântula (BEWLEY, 1994).

Rae *et al.* (2009) descreveram um rápido aumento na expressão de VPE entre o internó mais imaturo e o mais maduro. Também há relatos de que a  $\gamma$ VPE em *Arabidopsis* está envolvida na degradação da invertase vacuolar na senescência de tecidos e a mutação no gene da  $\gamma$ VPE levou a níveis maiores de glicosidase,  $\alpha$ -manosidases e uma  $\beta$ -galactosidase, sugerindo que essas proteínas podem ser processadas ou degradadas por VPEs (ROJO *et al.*, 2003). Um estudo da fitocistatina C-terminal estendida de *Arabidopsis thaliana* indica que essa cistatina tem um papel fundamental na germinação de sementes e crescimento de plântulas, e os autores sugerem que esse inibidor pode regular a atividade de cisteíno-peptidases endógenas do tipo papaína e legumaína (HWANG *et al.*, 2009). Portanto, visto que ocorreu a inibição *in vitro* da CaneLEG2 por CaneCPI-3, essa fitocistatina pode estar envolvida na regulação da atividade da legumaína, além de peptidases do tipo papaína durante o desenvolvimento da cana-de-açúcar. Resultados semelhantes foram encontrados por Zhao *et al.* (2013) que utilizaram a embriogênese de *Nicotiana tabacum* como um sistema modelo para demonstrar que o mecanismo de morte celular programada (MCP) no suspensor (órgão estrutural do embrião) é baseado na ação antagônica de duas proteínas: o inibidor de peptidase NtCYS e seu alvo, a peptidase do tipo catepsina H, NtCP14. Essa interação é fundamental para a ativação da MCP somente após o suspensor ter exercido sua função, levando ao desenvolvimento correto do embrião. Sem a ação do



inibidor, ocorre a morte celular precoce do suspensor e consequente aborto do embrião.

Estudos apontam que legumainas de plantas possuem atividade tipo caspase (ROJO *et al.*, 2004; HATSUGAI *et al.*, 2004; NAKAUNE *et al.*, 2005; BOSCH *et al.*, 2010; JULIÁN *et al.*, 2013). Baseando-se nesses estudos, o inibidor de caspase-1, Ac-YVAD-CHO foi testado contra a CaneLEG2 recombinante (Tabela 4.2, Figura 4.16 D). Foi observado que houve inibição, sugerindo que esta enzima pode estar relacionada à atividade do tipo caspase em cana-de-açúcar, mas uma análise mais específica é necessária para essa confirmação.

Apesar de não haver nenhuma caspase vegetal descrita, a utilização de substratos e inibidores dessa classe de enzimas tem indicado a presença de atividade relacionada à caspases em células vegetais (BOZHKOV *et al.*, 2002; ELBAZ *et al.*, 2002; BOSCH; FRANKLIN-TONG, 2007). Esse tipo de atividade em plantas está associado à morte celular programada (MCP), que ocorre durante o desenvolvimento, senescência de órgãos e resposta hipersensitiva (HR) após ataque de patógenos (PISZCZEK; GUTMAN, 2007). Para desvendar o mecanismo que constitui a base da morte celular mediada por VPEs em plantas, a maquinaria que desencadeia o colapso vacuolar precisa ser esclarecida. Como as VPEs agem por um processamento para ativar várias proteínas vacuolares, elas podem converter as enzimas hidrolíticas inativas em formas ativas, que então degradam os vacúolos e iniciam a cascata proteolítica da MCP (ISHIKAWA *et al.*, 2011). Entretanto, estudos mais aprofundados são necessários para demonstrar essa provável função em cana-de-açúcar.

## 5. CONCLUSÃO

- A busca e classificação das sequências do banco de dados SUCEST resultou na separação de três grupos distintos de legumainas de cana-de-açúcar.
- A análise *in silico* da CaneLEG2, representante do grupo II mostrou que ela possui os motivos conservados e resíduos catalíticos comuns em legumainas de plantas, além de um peptídeo sinal de 19 aminoácidos.
- O sistema de expressão heteróloga *P. pastoris* mostrou-se eficiente para a obtenção da CaneLEG2 recombinante.
- A indução da CaneLEG2 com um inibidor de cisteíno-peptidase, CaneCPI-3CT foi necessária para a purificação da enzima recombinante, evitando que ela sofresse processamento auto-catalítico precoce.
- A CaneLEG2 purificada apresentou atividade sobre os substratos comumente usados para legumainas, Z-Ala-Ala-Asn-MCA e Z-Val-Ala-Asn-MCA.
- A atividade enzimática da CaneLEG2 foi fortemente inibida *in vitro* pela CaneCPI-3 e CaneCPI-3CT, mas também apresentou inibição pela CaneCPI-3DeIC e pelo inibidor de caspase1 Ac-YVAD-CHO.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

ABE, K. *et al.* Molecular cloning of a cysteine proteinase inhibitor of Rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice seeds. **J Biol Chem**, v. 262, p. 16793-16797, 1987.

ABE, K. *et al.* Asparaginyl endopeptidase of jack bean seeds. Purification, characterization, and high utility in protein sequence analysis. **J Biol Chem**, v. 262, p. 16793-16797, 1992.

AKADA, R.; MURAKANE, T.; NISHIZAWA, Y. DNA extraction method for screening yeast clones by PCR. **Biotechniques**, v. 28, n. 4, p. 670-673, 2000.

AKAIKE, H. A new look at the statistical model identification. **Autom Control IEEE Trans**, v. 19, p. 716–723, 1974

ANNADANA, S. *et al.* Cloning, functional expression in *Pichia pastoris* and purification of potato cystatin and multicystatin. **J Biosci Bioeng**, v. 95, n. 2, p. 118-123, 2003.

ARAI, S. *et al.* Plant Seed Cystatins and Their Target Enzymes of Endogenous Origin. **J Agric Food Chem**. v. 50, p. 6612-6617, 2002.

ALIM, M. A. *et al.* Characterization of asparaginyl endopeptidase, legumain induced by blood feeding in the ixodid tick *Haemaphysalis longicornis*. **Insect Biochem Mol Biol**, v. 3, p. 911-922, 2007.

ALVAREZ-FERNANDEZ, M. *et al.* Inhibition of mammalian legumain by some cystatins is due a novel second reactive site. **J Biol Chem**, v. 274, n. 27, p. 19195-19203, 1999.

ALVAREZ-FERNANDEZ; ABRAHAMSON, M. Structural and protease inhibitory features of cystatins. In: ZEROVNIK, E.; JERALA, N. K. (eds) Human Stefins and Cystatins. Nova York: Nova Science Publishers, Inc., p. 23-42, 2006.

BARRETT, A. J. Proteases. **Curr Protoc Protein Sci**, sup. 21, p. 21.1.1-21.1.12, 2000.

BARRETT, A.J. *et al.* Proteinase Inhibitors. **Science Publishers**. v. 12, p. 515-569, 1986.

BENCHABANE, M. *et al.* Plant cystatins. **Biochimie**, v. 92, n. 11, p. 1657-1666, 2010.

BENDSTEN, J. D. *et al.* Improved prediction of signal peptides. SignalP 3.0. **J Mol Biol**, v. 340, p. 783-795, 2004.

BEYNON, R. J.; BOND, J.S. **Proteolytic enzymes**: a practical approach. Oxford University Press, 1990.

BEWLEY, J.D.; BLACK, M. **Seeds**: Physiology of Development and Germination, second ed. Plenum Press, New York, 1994.

BOSCH, M.; FRANKLIN-TONG, V.E. Temporal and spatial activation of caspase-like enzymes induced by self-incompatibility in *Papaver* pollen. **Proc Natl Acad Sci USA**, v. 104, n. 46, p. 18327-18332, 2007.

BOSCH, M. *et al.* Characterization of a legumain/vacuolar processing enzyme and YVADase activity in *Papaver* pollen. **Plant Mol Biol**, v. 74, n. 4-5, p. 381-393, 2010.

BOZHKOVA P.V., FILONOVA L.H., VON ARNOLD S. A key developmental switch during Norway spruce somatic embryogenesis is induced by withdrawal of growth regulators and is associated with cell death and extracellular acidification. **Biotechnol. Bioeng.** v. 77, n. 6 p. 658–667, 2002.

CARRILLO, L. *et al.* A barley cysteine-proteinase inhibitor reduces the performance of two aphid species in artificial diets and transgenic *Arabidopsis* plants. **Transgenic Res**, v. 20, n. 2, p. 305-319, 2011.

CHEN, J.M. *et al.* Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. **FEBS Lett**, v. 441, p. 361-365, 1998.

CHEN, H.J. *et al.* Molecular cloning and characterization of a cDNA encoding asparaginyl endopeptidases from sweet potato (*Ipomoea batatas* (L.) Lam) senescent leaves. **J Exp Bot**, v. 55, n. 398, p. 825-835, 2004.

CHENG, Y.; PRUSOFF, W. H. Relationship between the inhibition constant (K<sub>1</sub>) and the concentration of inhibitor which causes 50 per cent inhibition (I<sub>50</sub>) of an enzymatic reaction. **Biochemical Pharmacology**, v. 22, p. 3099-3108, 1973.

CHENG, T. *et al.* Cystatin M/E is a high affinity inhibitor of cathepsin V and cathepsin L by a reactive site that is distinct from the legumain-binding site. A novel clue for the role of cystatin M/E in epidermal cornification. **J Biol Chem**, v. 281, n. 23, p. 15893-15899, 2006.

CHU, M.H. *et al.* Crystal structure of tarocystatin-papain complex: implications for the inhibition property of group-2 phytocystatins. **Planta**, v. 234, n. 2, p. 243-54, 2011.

CONESA, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, v. 21, n. 18, p. 3674–3676, 2005.

COPELAND, R. A. **Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists.** New Jersey: John Wiley & Sons, Inc., 2005.

CORREA, G. C.; MARGIS-PINHEIRO, M.; MARGIS, R. Identification, classification and expression pattern analysis of sugarcane cysteine proteinases. **Genetics And Molecular Biology**, v. 24, n. 1, p.275-283, 2001.

CORRE-MENGUY, F. *et al.* Characterization of the expression of a wheat cystatin gene during caryopsis development. **Plant Mol Biol**, v. 50, n. 4-5, p. 687-698, 2002.

CREGG, J. M. *et al.* Recombinant protein expression in *Pichia Pastoris*. **Molecular Biotechnology**, v. 16, p. 23-52, 2000.

CREGG, J. M. **Methods in Molecular Biology: Pichia Protocols**, 2.ed., v. 389. Humana Press, Totowa, NJ, 2007.

DALL, E.; BRANDSTETTER, H. Mechanistic and structural studies on legumain explain its zymogenicity, distinct activation pathways, and regulation. **PNAS**, v. 110, n. 27, p. 10940-10945.

DILLEWIJN, C. V. Botany of Sugarcane, **Chronica Botanica**, Waltham, MA, 1952.

DIOP, N.N. *et al.* A multicystatin is induced by drought-stress in cowpea (*Vigna unguiculata* (L.) Walp. Leaves. **FEBS Lett**, v. 577, n. 3, p. 545-550, 2004.

ELBAZ, M.; AVNI, A.; WEIL, M. Constitutive caspase-like machinery executes programmed cell death in plant cells. **Cell death and differentiation**. v. 9, p. 726-733, 2002.

GANEVA, V.; GALUTZOV, B.; TEISSIÉ, J. Electric field mediated loading of macromolecules in intact yeast cells is critically controlled at the wall level. **Biochim Biophys Acta**, v. 1240, n. 2, p. 229-236, 1995.

GIANOTTI, A. *et al.* Recombinant expression, purification, and functional analysis of two novel cystatins from sugarcane (*Saccharum officinarum*). **Protein Expr Purif**, v. 47, p. 483,489, 2006.

GIRARD, C. *et al.* A multicomponent, elicitor-inducible cystatin complex in tomato, *Solanum lycopersicum*. **New Phytol**. V.173, p. 841-851, 2007.

GUINDON, S.; GASCUEL, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. **Syst Biol**, v. 52, p. 696–704, 2003.

GRUIS, D.F. *et al.* Redundant proteolytic mechanisms process seed storage proteins in the absence of seed-type members of the vacuolar processing enzyme family of cysteine proteases. **Plant Cell**, v. 15, n. 11, p. 2863-2882, 2002.

GRUIS, D.F.; SCHULZE, J.; JUNG, R. Storage protein accumulation in the absence of the vacuolar processing enzyme family of cysteine proteases. **Plant Cell**, v. 16, n. 1, p. 270-290, 2004.

HABIB, H.; FAZILI, K. M. Plant protease inhibitors: a defense strategy in plants. **Biotech Mol Biol Rev**, v. 2, n. 3, p. 68-85, 2007.

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

HARA-NISHIMURA, I.; HATSUGAI, N. The role of vacuole in plant cell death. **Cell Death Differ**, v. 18, p. 1298-1304, 2011.

HARA-NISHIMURA, I.; INOUE, K.; NISHIMURA, M. A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into mature forms. **FEBS Lett**, v. 294, p. 89-94, 1991.

HARA-NISHIMURA, I. *et al.* Vacuolar processing enzyme: an executor of plant cell death. **Current Opinion in Plant Biology**, v. 8, p. 404-408, 2005.

HATSUGAI, N. *et al.* A cellular suicide strategy of plants: vacuole-mediated cell death. **Apoptosis**, v. 11, p. 905-911, 2006.

HERMAN, E.M.; LARKINS, B.A.; Protein Storage Bodies and Vacuoles. **Plant Cell**, v. 11, p. 601-613, 1999.

HIRAIWA, N.; NISHIMURA, M.; HARA-NISHIMURA, I. Vacuolar processing enzyme is self-catalytically activated by sequential removal of the C-terminal and N-terminal propeptides. **FEBS Lett**, v. 447, p. 213-216, 1999.

HORTON, P. *et al.* WoLF PSORT: Protein Localization Predictor. **Nucleic Acids Res**, v. 35 (Web Server Issue), p. W585-7, 2007.

HWUANG, J. E. *et al.* Regulation of seed germination and seedling growth by an *Arabidopsis* phytocystatin isoform, AtCYS6. **Plant Cell Rep**, v. 28, n. 11, p. 1623-1632, 2009.

ISHIKAWA, T. *et al.* Bax inhibitor-1: a highly conserved endoplasmic reticulum-resident cell death suppressor. **Cell Death and Differentiation**, v. 18, p. 1271-1278, 2011.

IUBMB. Enzyme nomenclature recommendations. International Union of Biochemistry and Molecular Biology, Academic Press, Orlando FL, 1992.

JACKSON, M.A. *et al.* A bioinformatic approach to the identification of a conserved domain in a sugarcane legumain that directs GFP to the lytic vacuole. **Func Plant Biol**. v. 34, p. 633-644, 2007.

JULIÁN, I. *et al.* M. Phylogenetically distant barley legumains have a role in both seed and vegetative tissues. **Journal of Experimental Botany**, Vol. 64, No. 10, pp. 2929–2941, 2013.

KEMBHAVI, A. A. *et al.* The two cysteine endopeptidases of legume seeds: purification and characterization by use of specific fluorometric assays. **Arch Biochem Biophys**, v. 303, n. 2, p. 208-213, 1993.

KINOSHITA, T. *et al.* Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions. **Plant J**, v. 19, n. 1, p. 43-53, 1999.

KONDO, H. *et al.* Gene organization of oryzacystatin-II, a new cystatin superfamily member of plant origin, is closely related to that of oryzacystatin-I but different from those of animal cystatins. **FEBS Lett**, v. 278, n. 1, p. 87-90, 1991.

KUROYANAGI, M.; NISHIMURA, M.; HARA-NISHIMURA, I. Activation of Arabidopsis vacuolar processing enzyme by self-catalytically removal of an auto-inhibitory domain of the C-terminal propeptide. **Plant Cell Physiol**, v. 43, n. 2, p. 143-151, 2002.

LINNESTAD, C. *et al.* Nucellain, a barley homolog of the dicot vacuolar-processing protease, is localized in nucellar cell walls. **Plant Physiol**, v. 118, p. 1169-1180, 1998.

MARGIS, R.; REIS, E.M.; VILLERET, V. Structural and phylogenetic relationships among plant and animal cystatins. **Arch Biochem Biophys**, v. 359, n. 1, p. 24-30, 1998.

MARGIS-PINHEIRO, M. *et al.* Molecular evolution and diversification of plant cysteine proteinase inhibitors: new insights after the poplar genome. **Mol. Phylogenet. Evol.** v.49, p. 349-355, 2008.

MARTINEZ, M. *et al.* Comparative phylogenetic analysis of cystatin gene families from Arabidopsis, rice and barley. **Mol. Genet. Genomics** v.273, p. 423-432, 2005.

MARTINEZ, M. *et al.* Carboxy terminal extended phytocystatins are bifunctional inhibitors of papain and legumain cysteine proteinases. **FEBS Lett**, v. 581, n. 16, p. 2914-2918, 2007.

MARTINEZ, M. *et al.* 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, p. 1531–1545, 2009.

MOSOLOV, V.V.; VALUEVA, T.A. Participation of proteolytic enzymes in the interaction of plants and phytopathogenic microorganisms. **Biochemistry (Moscow)**, v. 71, n. 8, p. 838-845, 2006.

MÜNTZ, K.; SHUTOV, A.D. Legumains and their functions in plants. **Trends Plant Sci**, v. 7, n. 8, p. 340-344, 2002.

MÜNTZ, K.; BLATTNER, F.R.; SHUTOV, A.D. Legumains – a family of asparagine-specific cysteine endopeptidases involved in propolypeptide processing and protein breakdown in plants. **J Plant Physiol**, v. 159, p. 1281-1293, 2002.

NAKAUNE, S. *et al.* A vacuolar processing enzyme, deltaVPE, is involved in seed coat formation at the early stage of seed development. **Plant Cell**, v.17, n. 3, p. 876-887, 2005.

OKAMOTO, T.; MINAMIKAWA, T. Molecular cloning and characterization of *Vigna mungo* processing enzyme 1 (VmPE-1), an asparaginyl endopeptidase, **Plant Mol. Biol.** v. 39, p. 63-73, 1999.

OLIVA, M. L. *et al.* Inhibitory selectivity of canecystatin: a recombinant cysteine peptidase inhibitor from sugarcane. **Biochem Biophys Res Commun**, v. 320, n. 4, p. 1082-1086, 2004.

PERNAS, M.; SANCHEZ-MONGE, R.; SALCEDO, G. Biotic and abiotic stress can induce cystatin expression in chestnut. **FEBS Lett**, v. 467, n. 2-3, p. 206-210, 2000.

PISZCZEK, E.; GUTMAN, W. Caspase-like proteases and their role in programmed cell death in plants. **Acta Physiol Plant**, v. 29, p. 391-398, 2007.

RAE, A.L. *et al.* Functional specialization of vacuoles in sugarcane leaf and stem, **Trop. Plant Biol.** v. 2 , p. 13-22, 2009.

RAWLINGS, N. D.; MORTON, F. R.; BARRETT, A. J. An Introduction to Peptidases and the Merops Database. In: POLAINA, J.; MACCABE, A.P. (eds). **Industrial Enzymes**, p. 161-179, 2007.

RAWLINGS, N.D. *et al.* MEROPS: the peptidase database. **Nucleic Acids Res**, v. 36, p. 320-325, 2008.

RAWLINGS, N.D. *et al.* MEROPS: the peptidase database. **Nucleid Acids Res**, v. 42, p. D503 – D509, 2014.

RIVARD, D. *et al.* MsCYS1 a developmentally-regulated cystatin from alfalfa, **Plant Physiol. Biochem**, v. 45, p. 508-514, 2007.

ROJO, E. *et al.* VPEc exhibits a caspase-like activity that contributes to defense against pathogens. **Curr Biol** 14:1897–1906,2004.

ROTARI, V. I.; DANDO, P. M.; BARRET, A. J. Legumain forms from plants and animals differ in their specificity. **Biol Chem**, v. 382, p. 953-959, 2001.



SAMBROOK, J.; RUSSEL, D. W. **Molecular Cloning**: A laboratory Manual. 2.ed. New York: Cold Spring Harbor Laboratory Press, 2001.

SANTOS-SILVA, L. K. **Produção recombinante e caracterização funcional de uma legumaina de cana-de-açúcar**. 2012, 146 f. Tese (Doutorado em Genética Evolutiva e Biologia Molecular) – UFSCar, São Carlos, 2012.

SANTOS-SILVA, L. K. *et al.* Recombinant expression and biochemical characterization of sugarcane legumain. **Plant physiology and biochemistry**, v. 57, p. 181-192, 2012.

SCHECHTER, I.; BERGER, A. On the size of the active site in peptidases. I Papain. **Biochem Biophys Res Commun**, v. 27, p. 157-162, 1967.

SCHLERETH, A. *et al.* Comparison of globulin mobilization and cysteine proteinases in embryonic axes and cotyledons during germination and seedling growth of vetch (*Vicia sativa* L.). **J. Exp. Bot.** v. 51, p. 1423-1433, 2000.

SCOTT, M.P. *et al.* A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean. **Proc Natl Acad Sci USA**, v. 89, p. 658-662, 1992.

SHIMADA, T. *et al.* Vacuolar processing enzymes are essential for proper processing of seed storage proteins in *Arabidopsis thaliana*. **J Biol Chem**, v. 278, p. 32292-23399, 2003.

SIN, S. F.; YEUNG, E.C.; CHYE M.L. Downregulation of *Solanum americanum* genes encoding proteinase inhibitor II causes defective seed development, **Plant J.** v. 45, p. 58-70, 2006.

SOARES-COSTA, A. *et al.* A sugarcane cystatin: recombinant expression, purification and antifungal activity. **Biochem Biophys Res Commun**, v. 296, p. 1194-1999, 2002.

SOLOMON, M. *et al.* The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. **Plant Cell**, v. 11, n. 3, p. 431-444, 1999.

SUGAWARA, H. *et al.* Is a cysteine proteinase inhibitor involved in the regulation of petal wilting senescing carnation (*Dianthus caryophyllus*) flowers? **J Exp Bot**, v. 53, n. 368, p. 407-413, 2002.

TAMURA, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Mol Biol Evol**, v. 28, p. 2731–2739, 2011.

TIAN, L. *et al.* Differential proteomic analysis of soluble extracellular proteins reveals the cysteine protease and cystatin involved in suspension-cultured cell proliferation in rice. **Biochim Biophys Acta**, v. 1794, n. 3, p. 459-467, 2009.

TURK, V.; BODE, W. The cystatins: Protein inhibitors of cysteine proteinases. **FEBS Lett**, v. 285, n. 2, p. 213-219, 1991.

URWIN, P. E.; GREEN, J.; ATKINSON, H. J.; Expression of a plant cystatin confers partial resistance to Globodera, full resistance is achieved by pyramiding with natural resistance. **Mol Breeding**, v. 12, p. 263-269, 2003.

VETTORE, A. L. *et al.* Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome Research**, v. 13, n. 12, p. 2725-2735, 2003.

WANG, K.-M. *et al.* Characterization of inhibitory mechanism and antifungal activity between group-1 and group-2 phytocystatins from taro (*Colocasia esculenta*). **FEBS J.** v.275, p. 4980-4989, 2008.

WATERHAM, H. R. *et al.* Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. **Gene**, v.186, p.37-44, 1997.

WEEDA, S. M.; KUMAR, M. G. N.; KNOWLES, R. N. Developmentally linked changes in proteases and protease inhibitors suggest a role for potato multicystatin in regulating protein content of potato tubers. **Planta**, v. 230, n. 1, p. 73-84, 2009.

WOLTERING, E. J. Death proteases: alive and kicking. **Cell Press**, v. 15, n. 4, p. 185-188, 2010.

YAMADA, Y. *et al.* A cysteine protease from maize isolated in a complex with cystatin. **Plant Cell Physiol**, v. 41, n. 2, p. 185-191. 2000.

YANG, A. H.; YEH, K. W. Molecular cloning, recombinant gene expression, and antifungal activity of cystatin from taro (*Colocasia esculenta* cv. Kaosiung no.1). **Planta**, v. 221, n. 4, p. 493-501, 2005.

ZHANG, H.; ZHENG, X.; ZHANG, Z. The role of vacuolar processing enzymes in plant immunity. **Plant Signal Behav**, v. 12, p. 1565-1567, 2010.

ZHAO, Y. *et al.* Two wound-inducible soybean cysteine proteinase inhibitors have greater insect digestive proteinase inhibitory activities than a constitutive homolog. **Plant Physiol**, v. 111, n. 4, p. 1299-1306, 1996.

ZHAO P. *et al.* A Bipartite Molecular Module Controls Cell Death Activation in the Basal Cell Lineage of Plant Embryos. **PLoS Biol** v.11, n. 9, 2013.

ZHAO, L. *et al.* Structural analysis of asparaginyl endopeptidase reveals the activation mechanism and a reversible intermediate maturation stage. **Cell Res** v. 24, p. 344-358, 2014.