

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**“CARACTERIZAÇÃO FUNCIONAL DE UMA CISTEÍNO
CATEPSINA RECOMBINANTE DA FORMIGA CORTADEIRA
ATTA SEXDENS”**

Katia Celina Santos Correa*

Dissertação apresentada como parte dos requisitos para obtenção do título de MESTRE EM QUÍMICA, área de concentração: QUÍMICA.

Orientador (a): Prof.(a) Dr.(a) Dulce Helena Ferreira de Souza

*** bolsista (FAPESP)**

**São Carlos - SP
2018**



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Química

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata Katia Celina Santos Corrêa, realizada em 27/04/2018:

Profa. Dra. Dulce Helena Ferreira de Souza
UFSCar

Prof. Dr. Flavio Henrique da Silva
UFSCar

Prof. Dr. Julio Cesar Borges
IQSC/USP

Ao José Correa por todo o incentivo.

*“...são as águas de Março fechando verão,
é promessa de vida no teu coração...”.*

Tom Jobim

AGRADECIMENTOS

À Profa. Dra. Dulce Helena Ferreira de Souza, pela orientação, confiança e ensinamentos.

À Profa. Dra. Maria Teresa M. Novo Mansur por permitir que parte do trabalho fosse feito em seu laboratório. Aos seus alunos pelo carinho e atenção, em especial para o Vinícius pelas discussões e ensinamentos.

Ao Prof. Dr. Paulo Cesar Vieira por permitir o uso do equipamento para que os experimentos da atividade enzimática fossem realizados.

Ao Prof. Dr. Flávio Henrique-Silva pela colaboração, cedendo os inibidores proteicos para que os ensaios de inibição fossem realizados, um divisor de águas durante a realização desse projeto.

À Dra. Ariele Cristina Moreira pela colaboração, pelos cDNAs da formiga *A. sexdens*.

À Dra. Cláudia Alves pelas discussões, dicas e atenção.

Ao Dr. Hugo César Ramos de Jesus pela colaboração com os experimentos da espectrometria de massas.

Ao aluno de doutorado Amr Galal pela ajuda com o programa estatístico R.

Aos ex e atuais integrantes do LBFE - Laboratório de Bioquímica Funcional e Estrutural, Júlio, Caroides, Guilherme, Caio, Mônica, Priscila, Ariele, Mariana, Tatiane, Renata, Camila, Cláudia, Paulo, Bruno, Bruna, Adriana, Evandro, Melissa, Juliana, Kelly e Hugo, pelas contribuições e convivência.

Aos amigos do coração Júlio (*in memoriam*), Jhony, Fábria, Jocely, Luane, Déborah, Bruna, Diego, Caroides, obrigada pelo companheirismo e pela alegria de ter vocês sempre por perto.

Às secretárias da Pós-Graduação Ariane, Luciani e Cristina por toda atenção.

Ao Programa de Pós-Graduação em Química da Universidade Federal de São Carlos - UFSCar.

À Fundação de Amparo à Pesquisa do Estado de São Paulo FAPESP, processo 2015/21517-7, pelo suporte financeiro.

LISTA DE ABREVIATURAS

- cDNA: DNA complementar a um mRNA
- DNA: Ácido desoxirribonucleico (*Deoxiribonucleic acid*)
- dNTP: Desoxinucleosídeos trifosfatos
- E. coli*: *Escherichia coli*
- EDTA: Ácido etilenodiamino tetracético (*Ethylenediamine Tetraacetic Acid*)
- g: Unidade de força rotocentrífuga (gravidade)
- E-64: L-trans-Epoxysuccinyl leucylamido (4-guanidino) butane
- kb: Kilobase
- kDa: Kilodalton(s), unidade de massa molecular
- LC: Cromatografia Líquida
- mg: Miligrama
- mL: Mililitro
- mM: Milimolar
- mRNA: Ácido ribonucleico mensageiro
- MS/MS: Espectrometria de massas sequencial
- ng: Nanograma
- nm: nanômetro (unidade de medida correspondente a 10^{-9} metros)
- Ni-NTA: Ácido Nitriliacético carregado com íons níquel
- ORF: Fase de leitura aberta (*Open Reading Frame*)
- pb: Pares de bases
- pH: Potencial de hidrogênio
- SDS: Dodecil sulfato de sódio (*Sodium Dodecyl Sulfate*)
- SDS-PAGE: eletroforese em gel de poliacrilamida e SDS (*Polyacrylamide Gel Electrophoresis*)

LISTA DE TABELAS

TABELA 3.1 - Condições dos ciclos da reação de PCR para amplificação do gene.	23
TABELA 3.2 - Condições da reação de amplificação do DNA	23
TABELA 3.3- <i>Condições da reação de digestão EcoRI e NcoI</i>	27
TABELA 3.4 - Condições da reação de ligação	28
Tabela 3.5 - Condições da reação de amplificação do clone.	29

LISTA DE FIGURAS

FIGURA 1.1- Operárias de <i>Atta colombica</i> carregam incessantemente folhas para o ninho para usar como substrato para o cultivo do fungo mutualista.....	2
FIGURA 1.2 - Relação simbiótica entre o fungo <i>Leucocoprinus gongylophorus</i> e a formiga “cortadeira”	2
FIGURA 1.3- Relação simbiótica entre os organismos no interior do ninho da formiga	4
FIGURA 1.4 - Árvores de <i>Eucalyptus grandis</i> desfolhadas aos 6 meses de idade	5
FIGURA 1.5 - Esquema representativo dos mecanismos de ativação das peptidases	7
FIGURA 1.6 - Intermediários tetraédricos formados durante a clivagem peptídica.....	9
FIGURA 1.7 - Estrutura cristalográfica da pró-catepsina L e da catepsina L, enzima digestiva da larva do inseto <i>Tenebrio molitor</i> , denominada pCAL3C26S.....	12
FIGURA 1.8 - Mecanismo da hidrólise do substrato pela enzima cisteíno peptidase <i>papain-like</i>	13
FIGURA 1.9 - Interação dos subsítios da enzima com os resíduos de aminoácidos do substrato.....	14
FIGURA 1.10 - Representação em 3D dos subsítios de ligação da enzima com o substrato ao longo da cavidade catalítica	14
FIGURA 1.11 – Modelo estrutural para a cistatina de planta, orizacistatina. Modelo construído por modelagem molecular	17
FIGURA 3.1- Sequência de nucleotídeos	22
Figura 3.2 - Sequência do sítio de clonagem múltipla dos vetores pET-32a-c(+)	27
FIGURA 3.3 - Esquema geral dos ensaios enzimáticos na presença do substrato fluorogênico Z-FR-MCA.....	36
FIGURA 4.1- Tradução da sequência completa de DNA (completo CDS) em sequência de aminoácidos da catepsina L de <i>A.echinatior</i>	42
FIGURA 4.2- Oligonucleotídeos sintetizados para a amplificação da ORF CathL	43
FIGURA 4.4- Alinhamento das sequências de nucleotídeos de uma catesina da <i>A. echinatior</i> com a sequência do gene obtido CathL. A sequência sombreada em cinza é a região do peptídeo sinal da <i>A. echinatior</i> , ausente na sequência predita do CathL. (*) resíduos idênticos em ambas às sequências.....	44
FIGURA 4.5 - Análise qualitativa da expressão da ORF CathL em gel de agarose 1%	45

FIGURA 4.6 - Análise em gel de agarose 1% do DNA genômico extraído da cabeça da formiga <i>A. sexdens</i> e amplificação da ORF a partir do DNA genômico	47
FIGURA 4.8- Análise em gel de agarose 1% da análise de restrição utilizando as enzimas de restrição <i>NcoI</i> e <i>EcoRI</i>	47
FIGURA 4.9- Alinhamento das sequências preditas para catepsinas L da vespa <i>D. coarctata</i> , da catepsina L <i>A. Sexdens</i> (trabalho), da formiga <i>A. cephalotes</i> e de <i>A. echinator</i>	49
FIGURA 4.10 - Sequência de aminoácidos da AsCathL obtida pela tradução da sequência de nucleotídeos.....	50
FIGURA 4.11- Análise da sequência e classificação	51
FIGURA 4.12 - A - Análise em SDS-PAGE. B - Western blot (anti-His)	52
FIGURA 4.13 - A - Perfil cromatográfico. B - Análise em SDS-PAGE	54
FIGURA 4.14- Análise em SDS-PAGE da proteína recuperada a partir dos corpos de inclusão, utilizando o <i>refolding</i>	56
FIGURA 4.15 - Análise em gel SDS PAGE da proteína incubada em pH ácido, variando tempo de incubação e reconhecimento do anticorpo anti-His	57
FIGURA 4.16- Atividade da AsCathL dependente do pH usando o Z-Phe-Arg-AMC como substrato.....	58
4.17 - Efeito do tempo de pré-incubação em pH 4,5 na atividade de AsCathL recombinante. Atividade hidrolítica da AsCathL contra o substrato fluorgênico no intervalo de 30 minutos	59
FIGURA 4.18- Curva para determinação da constante de Michaelis-Menten (K_m) do substrato Z-Phe-Arg-MCA frente à catepsina L (AsCathL)	60
FIGURA 4.19 - Avaliação da interferência do epóxido E-64 e das cistatinas recombinantes de cana-de-açúcar (CaneCPI-1, CaneCPI-2, CaneCPI-3, CaneCPI-4, CaneCPI-5, CaneCPI-6) na atividade da catepsina L (AsCathL)	61
FIGURA 4.20 - Efeito das canacistatinas na atividade da AsCathL	63
FIGURA 4.21 - sequência de nucleotídeos da construção da AsCathL. Em vermelho, os peptídeos trípticos identificados por espectrometria de massas.....	64
Figura 4.22- Análise filogenética entre as catepsinas L das formigas depositadas no banco de dados NCBI e a predita catepsina da formiga <i>A. sexdens</i>	65
Figura 4.23- Porcentagem de identidade entre as sequências estudadas	66
Figura 4.24- Porcentagem de identidade	67

Figura 4.25 - Análise filogenética entre as catepsinas L dos insetos depositadas no banco de dados NCBI e a predita catepsina da formiga *A. sexdens*.....68

RESUMO

CARACTERIZAÇÃO FUNCIONAL DE UMA CISTEÍNO CATEPSINA RECOMBINANTE DA FORMIGA CORTADEIRA *ATTA SEXDENS*. As formigas cortadeiras, por dependerem de folhas vegetais para sobreviver, apresentam alto poder de prejuízo para plantações agrícolas, florestas naturais e plantadas. O controle desses insetos pelo uso de inseticidas comerciais além de não seletivo, é tóxico. A catepsina L, enzima da família das cisteíno peptidases, é conhecida por ser ativa na degradação da vitelina embrionária, no remodelamento dos tecidos, nos processos reprodutivos e na degradação da cutícula do inseto durante a muda, o que a torna um excelente alvo de estudo para busca de promissores inibidores. A sequência de uma catepsina de *Acromyrmex echinator* foi usada para desenhar *primers* que, juntamente com o cDNA de *A. sexdens*, foram utilizados para amplificação da ORF codificante para uma catepsina L da formiga *A. sexdens* utilizando PCR. A ORF denominada CathL, é expressa nos estágios de desenvolvimento: larva, pupa, e adulto, como também em partes do inseto adulto: cabeça, tórax e gaster. O gene foi amplificado quando se utilizou DNA genômico extraído da cabeça da formiga e utilizado como molécula molde no PCR, sugerindo que a ORF é oriunda da formiga e não do fungo *L. gongylophorus* com quem vive em simbiose. A ORF foi clonada em vetor pET32a e a enzima recombinante AsCathL foi expressa em *E.coli* majoritariamente como corpos de inclusão, e por *refolding*, foi obtida a pró-enzima purificada de massa molecular aparente de 52 kDa com rendimento de 5 mg/L de proteína por meio de cultura. A ativação da proteína foi feita pela remoção do pró-peptídeo N- terminal da pró-enzima AsCathL sobre condições ácidas e a enzima mostrou estar ativa, hidrolisando o substrato fluorogênico Z-Phe-Arg-AMC. A enzima mostrou maior atividade proteolítica em pH 4,5. Ensaio de inibição da atividade enzimática foram realizados com seis canacistatinas recombinantes de cana-de-açúcar, as quais demonstraram ser excelentes inibidores da catepsina, com valores de K_i que variam de 2,95 nM a 0,6 nM. Os resultados obtidos abrem perspectivas para busca de mais conhecimentos sobre essa cisteíno peptidase, um importante alvo para estudos no controle do inseto.

ABSTRACT

FUNCTIONAL CHARACTERIZATION OF A RECOMBINANT CATHEPSIN LIKE CYSTEINE FROM LEAF-CUTTING ANT *ATTA SEXDENS*. Cutting ants depend on plant leaves to survive, presenting high damage power in agricultural productions, natural and planted forest. The control of these insects through the use of commercial insecticides besides being toxic, it is not selective. Cathepsin L, an enzyme of the cysteine peptidase family, is known to be active in the degradation of embryonic vitellin, in the tissue remodeling, in the reproductive processes and in the cuticle degradation during insect moulting, which make this enzyme an excellent target for the study of promising inhibitors. The cathepsin sequence from *Acromyrmex echinator* was used to design primers, that together cDNA of the *A. sexdens*, were used for amplification of the ORF encoding a cathepsin L of the ant *A. sexdens* by PCR. The ORF called CathL is expressed in the ant developmental stages (larvae, pupae, adult), and is present in parts of the adult insect (head, mesosoma, gaster). The gene was amplified using genomic DNA extracted from the ant's head and used as a template molecule in the PCR, suggesting that the ORF comes from the ant and not from the fungus *L. gongylophorus* with whom it lives in symbiosis. The ORF was cloned into vector pET32a and the recombinant enzyme AsCathL was expressed in *E. coli* mainly as inclusion bodies, and through refolding, a purified proenzyme of 52 kDa was obtained with final yield of 5 mg/L. Activation of the protein allowed the removal of the N-terminal pro-enzyme AsCathL on acidic conditions and showed hydrolytic activity *in vitro* towards the synthetic substrate Z-Phe-Arg-AMC. The enzyme showed higher proteinase activity at pH 4.5. Inhibition assays of the enzymatic activity were carried out using six recombinant sugarcane canacistatins, which demonstrated to be excellent cathepsin inhibitors, with K_i ranging from 2,95 nM to 0,6 nM. The obtained results show perspectives to the pursuit of more knowledge about that cysteine peptidase, an important objective for the studies on the insect control.

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Formigas cortadeiras	1
1.2	Peptidases	6
1.3	Cisteíno peptidases.....	9
1.3.1	Inibidores das cisteíno peptidases	15
2	OBJETIVOS	20
2.1	Objetivo Geral	20
	Os objetivos deste trabalho foram a obtenção de uma catepsina L recombinante de <i>A. sexdens</i> (formiga saúva), a caracterização bioquímica e cinética e ensaios de inibição <i>in vitro</i> por canacistatinas.....	20
2.2	Etapas para alcançar os objetivos.....	20
3	METODOLOGIA.....	21
3.1	Extração de RNA de <i>A. sexdens</i> e síntese de cDNA	21
3.2	Clonagem da ORF	22
3.2.1	Amplificação do DNA que codifica a síntese da catepsina L ..	22
3.2.2	Análise qualitativa da expressão da ORF	24
3.2.3	Extração do DNA genômico.....	24
3.2.4	Digestão do vetor pET32a-c(+) e do produto de PCR.....	26
3.3	Expressão recombinante da AsCathL em <i>E.coli ArcticExpress</i> ..	30
3.4	Análise da expressão da pró-enzima por <i>Western blot</i>	31
3.5	Purificação da pró-enzima.....	32
3.6	Recuperação da pró-enzima a partir de corpos de inclusão.....	33
3.7	Ensaio enzimático para determinação da atividade proteolítica da cisteíno peptidase.....	34
3.7.1	Ativação da cisteíno peptidase	34
3.7.2	Caracterização cinética	36
3.8	Ensaio de inibição da atividade proteolítica	36

3.9	Análise da AsCathL por espectrometria de massas	38
3.10	Análise das sequências, alinhamentos e filogenia	40
4	RESULTADOS E DISCUSSÃO.....	42
4.1	Amplificação do DNA referente à catepsina de <i>A. sexdens</i>	42
4.2	Clonagem da ORF (CathL) em vetor de expressão e análise da sequência	47
4.3	Expressão da proteína recombinante em <i>E.coli</i>	51
4.4	Recuperação da AsCathL a partir dos corpos de inclusão	55
4.5	Caracterização cinética da enzima AsCathL.....	58
4.6	Inibição da atividade da enzimática da AsCathL	61
4.7	Identificação dos aminoácidos da catepsina por LC/MS.....	63
4.8	Análise dos dendrogramas	64
5	CONSIDERAÇÕES FINAIS E PERSPECTIVAS.....	69
6	REFERÊNCIAS BIBLIOGRÁFICAS	71
	ANEXO .A	83
	ANEXO .B	87

1 INTRODUÇÃO

1.1 Formigas cortadeiras

A humanidade vive em um planeta rodeado por insetos, dentre eles, os cupins, besouros, moscas, abelhas, vespas e formigas (WILSON e HOLLDOBLER, 2005). As formigas são notáveis entre os artrópodes e representam a metade da biomassa global de insetos. Dentre as milhares de espécies de formigas que estão catalogadas, destacam-se as “cortadeiras” de folhas, com uma elevada porcentagem na sociedade das formigas (BENCKISER, 2010; BOULOGNE et al., 2012). A dominância ecológica e o comportamento social dessas formigas têm atraído a atenção dos pesquisadores. Em termos de diversidade de espécies, abundância relativa, impacto ecológico e hábitos sociais, as formigas emergem como um dos grupos mais proeminentes dos artrópodes (WARD, 2006).

No século XIX, em uma pesquisa feita na floresta tropical brasileira, foi apurado que a quantidade de formigas naquele ambiente era de aproximadamente quatro vezes mais que a biomassa de todos os vertebrados presentes (WILSON, 1990), dados que ainda são relevantes segundo estudos publicados no ano de 2016 pelo Instituto de Pesquisas e Estudos Florestais (IPEF), responsável pela criação de técnicas para a melhoria da qualidade e produtividade das florestas no Brasil.

De acordo com o IPEF, dentre as mais de mil espécies de formigas existentes no Brasil, as saúvas e quenquéns conhecidas como “cortadeiras”, somam juntas cerca de 40 espécies e são consideradas nativas do país. Elas estão inseridas como insetos sociais que apresentam castas e se desenvolvem por holometabolismo (ovo-larva-pupa-adulto) (ZANETTI et al., 2002).

Pertencem à tribo *Attini*, são constituídas por 14 gêneros (ARAÚJO et al., 2015). Destes 14 gêneros, *Atta* (saúvas) e *Acromyrmex* (quenquéns) têm sido as mais estudadas. As formigas da tribo *Attini* vivem em simbiose com o fungo *Leucoagaricus gongylophorus*, pertencente à classe dos Basideomicetos (HERVEY et al., 1977), o qual proporciona às formigas nutrientes (MARTIN et al., 1969) e enzimas (BOYD e MARTIN, 1975) e em troca, as formigas proporcionam ao fungo uma variedade de substratos, estimulando o crescimento do fungo (WEBER, 1972; MARTIN et al., 1975). As formigas transportam o vegetal para o ninho, conforme

ilustrado na Figura 1.1 e o fungo cresce sobre esse vegetal conhecido como “Jardim de fungo” (WEBER, 1995), uma estrutura esponjosa de cor esbranquiçada, conforme aparência exibida na Figura 1.2.

FIGURA 1.1- Operárias de *Atta colombica* carregam incessantemente folhas para o ninho para usar como substrato para o cultivo do fungo mutualista



Fonte: WENSELEERS (2009)

FIGURA 1.2 - Relação simbiótica entre o fungo *Leucocoprinus gongylophorus* e a formiga “cortadeira”



Fonte: BOULOGNE et al., (2014)

Segundo MARTIN e WEBER (1969), nessa relação, o fungo produz a celulase, enzima responsável pela degradação da celulose gerando nutrientes para as formigas. Essa teoria foi reforçada pelas descobertas feitas por (BACCI et al., 1995) que mostraram que o *L. gongylophorus* era capaz de degradar a celulose. No entanto, estudos adicionais indicaram que a degradação da celulose era menor quando comparada à degradação de outros polissacarídeos presentes no material vegetal, como o xilano, o amido e a pectina. Uma vez que o fungo degrada e assimila esses polissacarídeos, ele é capaz de mediar a transferência de nutrientes

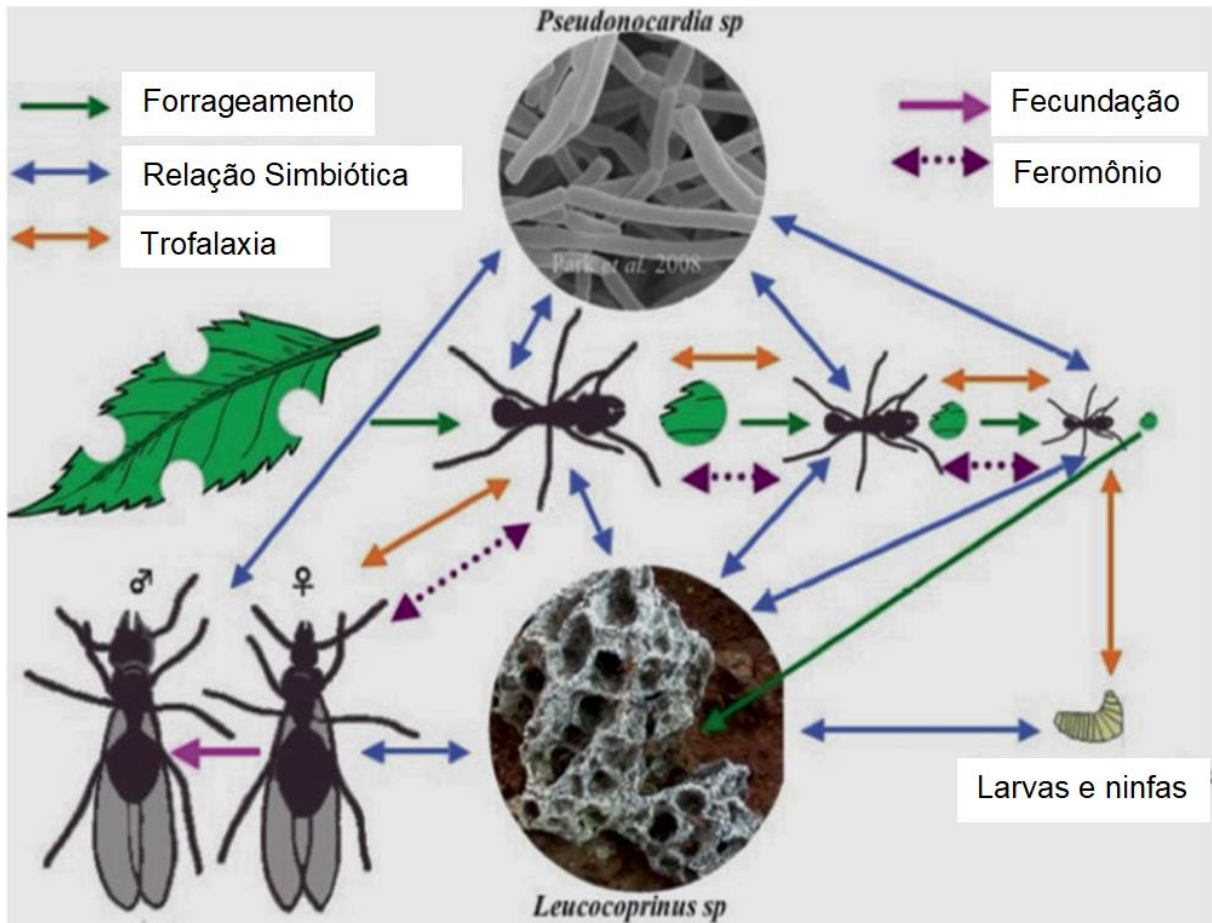
assimiláveis pelas formigas, como é o caso da glicose, a partir da degradação do amido. Essa integração metabólica aparenta ser uma parte importante na relação simbiótica fungo-formiga (SIQUEIRA et al., 1998; SILVA et al., 2003).

O fungo é a fonte de alimento para larvas, ninfas, rainha e complementam a dieta vegetal dos trabalhadores adultos (BOULOGNE et al., 2012). Em retorno, as formigas protegem o fungo por meio da *Pseudonocardia sp*, espécie de bactéria que vive nas cutículas das formigas. Essa relação simbiótica começa poucos dias após a formiga se tornar adulto e atinge a sua maior cobertura corporal de 13 a 18 dias. Essa bactéria possui propriedades antibióticas e antifúngicas que inibem o crescimento do fungo *Scovopsis sp*, aniquilador do fungo *L. gongylophorus* (CAMARGO et al., 2006; VIEIRA et al., 2017). Além dessa proteção, as formigas conseguem proporcionar ao fungo um ambiente livre de vários outros micro-organismos que possam interferir no desenvolvimento e manutenção do “jardim de fungo”, graças à ação de antibióticos que são produzidos pelas formigas e excretados pela glândula metapleurálica presente no organismo do inseto (POULSEN et al., 2003; VIEIRA et al., 2010).

Outra importante contribuição para a relação simbiótica é que, após a mastigação do vegetal, a formiga deposita um líquido fecal sobre o substrato. As enzimas proteolíticas presentes nesse material proporcionam uma melhor degradação da planta, favorecendo a colonização pelo fungo (BOYD e MARTIN, 1975; CAMARGO et al., 2006).

A relação mutualista é seguida pelo sistema conhecido como *foraging* (forrageamento) (FOWLER, 1978; CAMARGO et al., 2006), condicionado por uma trilha de feromônio que ocorre ao longo de uma linha envolvendo todas as castas da colônia. As formigas operárias, após descobrirem uma fonte vegetal, limpam o percurso a percorrer, permitindo que as formigas cheguem de forma mais rápida, acelerando a agregação da massa operária na trilha de feromônio. O comprimento da trilha de forrageamento chega até 100 metros do ninho (FOWLER, 1978; BOULOGNE et al., 2014). Assim, as formigas de forma organizada e sincronizada, levam ao desfolhamento de pequenas, médias e grandes áreas vegetais. A Figura 1.3 ilustra as complexas relações do mutualismo entre o ninho da formiga *Acromyrmex octospinosus*, a simbiose do *L. gongylophorus* e o simbiótico *Pseudonocardia sp*, assim como a ação do complexo sistema de *foraging* sobre a vida social da formiga cortadeira.

FIGURA 1.3- Relação simbiótica entre os organismos no interior do ninho da formiga



Fonte: Adaptada de BOULOGNE et al., (2014)

Essa espécie, por depender de folhas vegetais frescas para sobreviver, causa prejuízos para plantações agrícolas, florestas naturais e plantadas. Estima-se que as cortadeiras causem sérios danos na produção de algumas culturas, levando a perdas econômicas consideráveis (BOULOGNE et al., 2014). O ataque às plantas por esses insetos afeta de forma indireta a produção, pois com o corte, as plantas ficam menos resistentes e mais susceptíveis a ataques por outros insetos e patógenos (CANTARELLI et al., 2008).

Esse inseto pode causar a desfolhação total das plantas em espécies de eucaliptos, afetando tanto o diâmetro como a altura das árvores. Uma única colônia de formiga cortadeira por hectare de floresta pode reduzir o crescimento anual da plantação de *Eucalyptus* em 5% e de *Pinus* em 10% (ZANETTI et al., 2003; ZANETTI et al., 2014). Esses números exemplificam o prejuízo econômico que a presença desses organismos fitófagos acarreta, levando a uma diminuição da

produção e conseqüente redução de lucro. A Figura 1.4 ilustra o poder de destruição das formigas cortadeiras.

O combate a esses insetos, considerados pragas agrícolas, é uma preocupação constante e vários métodos têm sido utilizados como tentativas de controle. O método químico com base em inseticidas é o mais usual, porém eles são tóxicos e têm suas aplicações restritas pelo protocolo de Kyoto (LIVRAMENTO et al., 2018) .

FIGURA 1.4 - Árvores de *Eucalyptus grandis* desfolhadas aos 6 meses de idade



Fonte: Unibrás Agroquímica

Embora as formigas estejam vinculadas na maioria dos estudos e pesquisas, aos danos econômicos e ambientais causados, há de se considerar a importância dessas formigas ao meio ambiente. Esses insetos contribuem para a construção de um ambiente propício ao crescimento de determinadas plantações que se beneficiam do solo aerado (devido à construção dos ninhos) e da matéria orgânica (resíduo dos ninhos não aproveitados pelos organismos que ali habitam) (SUEN et al., 2011).

Estudar enzimas que são importantes para a sobrevivência desse inseto pode ser uma estratégia para seu controle. Nesse âmbito, as cisteíno peptidases têm sido relatadas como enzimas responsáveis tanto pelo desenvolvimento como pela sobrevivência da classe dos insetos. Diversos estudos apontam o envolvimento das catepsinas L de insetos na degradação da vitelina

embrionária (CHO et al., 1999), muda e metamorfose (TAKAHASHI et al., 1993; HEGEDUS et al., 2002), no remodelamento de tecidos (HOMMA et al., 1994), na embriogênese (YAMAMOTO e TAKAHASHI, 1993), nos processos reprodutivos (MATSUMOTO et al., 1997) e na degradação da cutícula do inseto durante a muda (LIU et al., 2006).

Em contraste com animais maiores, as cisteíno peptidases em invertebrados, em especial nos insetos, também são descritas por atuarem como enzimas digestivas (TERRA e FERREIRA, 1996; SOARES-COSTA et al., 2012; BETON et al., 2012). E apesar de estarem presentes no intestino do organismo da *Delia radicum* (HEGEDUS et al., 2002) e da mosca-da-fruta (*Drosophila melanogaster*) (MATSUMOTO et al., 1995), as cisteíno peptidase estão presentes também em outros tecidos desses organismos e não foram consideradas enzimas digestivas. Também há pesquisas que mostram atividade inseticida da catepsina L da mosca de carne (*Sarcophaga peregrina*) contra mariposas de tomate (*Lacanobia oleracea*) (PHILIP et al., 2007) e mariposas de flores (*Heliothis virescens*) (LI et al., 2008).

Segundo TERRA e FERREIRA (2005), as catepsinas L são as únicas cisteíno peptidases de insetos quantitativamente importantes, principalmente sob o ponto vista do desenvolvimento de técnicas de controle dos insetos considerados pragas, dado a importância que as catepsinas possuem tanto no desenvolvimento como na sobrevivência dos insetos. Portanto, são consideradas um excelente alvo de estudo para busca de inibidores.

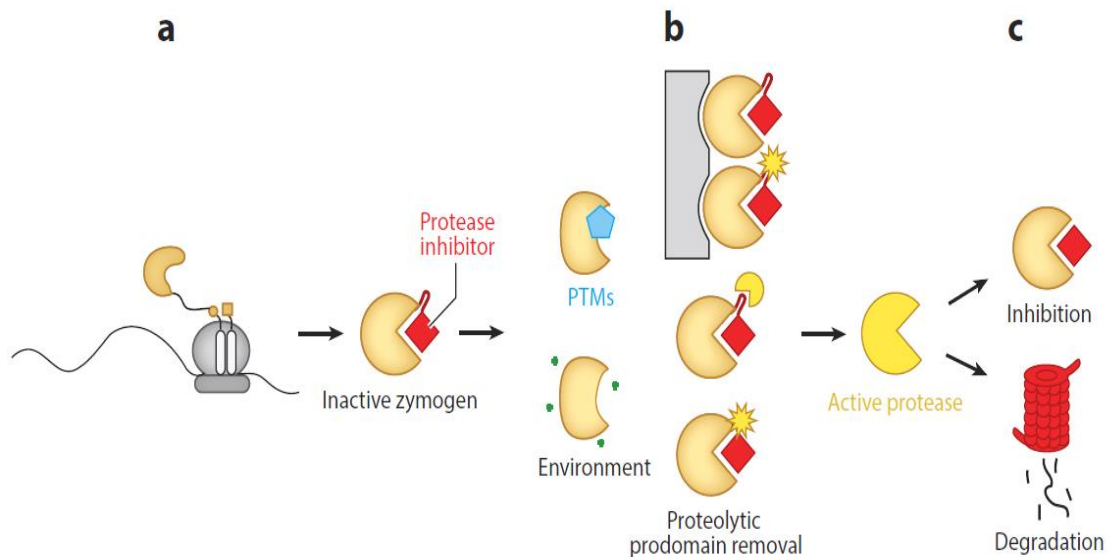
1.2 Peptidases

As peptidases ou enzimas proteolíticas são eficientes executores de uma reação química: a hidrólise das ligações peptídicas. Essas enzimas são componentes essenciais de muitos processos celulares e extracelulares, desde a maturação das proteínas até a morte celular em todos os organismos vivos. No entanto, essas enzimas catalisam reações hidrolíticas irreversíveis e, conseqüentemente, devem ser estritamente reguladas. A ação das peptidases pode ser controlada *in vivo* por vários mecanismos: regulação da expressão gênica, ativação de seus zimogênios inativos, bloqueio por inibidores endógenos, direcionamento para compartimentos específicos: lisossomos e mitocôndrias e

através de modificações pós-traducionais tais como glicosilação, formação de ligações dissulfeto, ligação metálica, proteólise e degradação (BERTENSHAW et al., 2003; LÓPEZ-OTÍN e BOND, 2008).

As peptidases são uma subclasse da classe das hidrolases (*Enzyme Commission* 3.4). A maioria dessas enzimas é sintetizada na forma de zimogênio, um precursor enzimático inativo que contém um peptídeo chamado pró-domínio o qual deve ser removido para ativar a enzima madura. Peptidases que são sintetizadas com o pró-domínio requerem ligações com cofatores ou modificações pós-traducionais para serem ativadas. A Figura 1.5 ilustra a esquematização da síntese de uma peptidase na forma de zimogênio e os mecanismos de ativação da mesma por sinalização macromolecular (*post translational modifications*- PTMs), por ação de outra protease ou por auto-proteólise: **a)** Peptidase sintetizada na forma de zimogênio no ribossomo; quadrado vermelho representa o pró-domínio; **b)** mecanismos de remoção do pró-domínio e **c)** após a ativação, a atividade da peptidase pode ser limitada por inibidores endógenos ou por degradação no proteassomos.

FIGURA 1.5 - Esquema representativo dos mecanismos de ativação das peptidases



Fonte: Adaptada de SANMAN e BOGYO (2014)

Baseado nos diferentes mecanismos catalíticos, as peptidases foram classificadas em endopeptidases, enzimas que clivam as ligações peptídicas ao longo da cadeia polipeptídica e em exopeptidases (aminopeptidases e

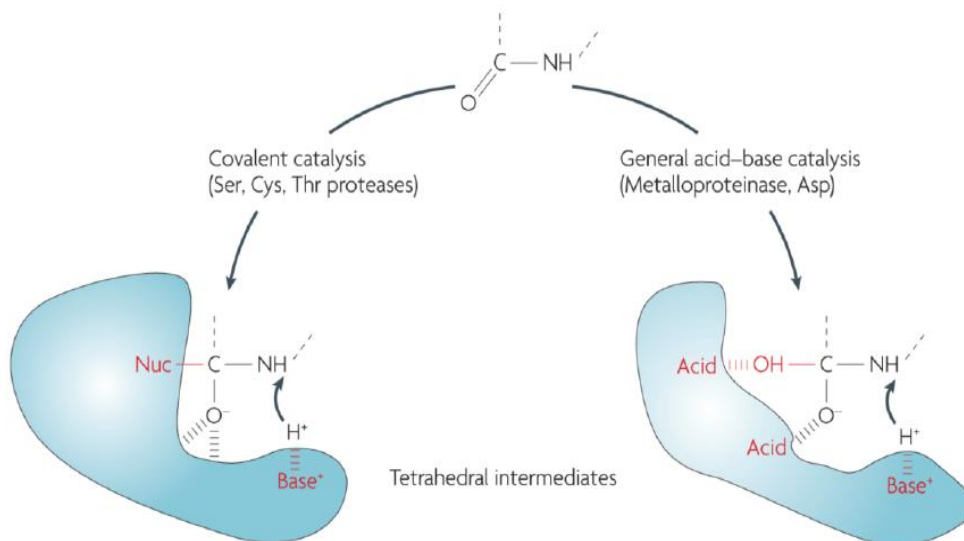
carboxipetidases), cuja ação proteolítica é direcionada para a região próxima ao agrupamento amina (NH_2) ou ao grupo carboxil (COOH). No entanto, a disponibilidade de informações estruturais e de mecanismos de ação destas enzimas levou a uma nova classificação, e em adição aos diferentes mecanismos catalíticos, as peptidases foram classificadas em seis subclasses. Dentre as mais conhecidas estão: serino (EC 3.4.21), cisteíno (E.C. 3.4.22), aspartil (E.C 3.4.23), metalo (E.C 3.4.24), e as menos conhecidas treonino (3.4.25) e a glutamil peptidases (DRAG e SALVESEN, 2010; TURK et al., 2012a).

Em cada subclasse mencionada, as peptidases diferem em especificidades em relação ao substrato pelo fato de cada enzima conter em seu sítio ativo aminoácidos específicos. As metalo peptidases necessitam de um íon metálico, as aspartil peptidases são ativas em pH ácidos e o mecanismo da catálise é baseado em dois resíduos de aspartato, as cisteíno peptidases possuem um resíduo de cisteína, de histidina e de asparagina e as serino peptidases, um resíduo de serina, de histidina e de ácido aspártico no sítio catalítico (TERRA et al, 2005).

As peptidases aceleram a formação de um estado de transição na hidrólise da ligação peptídica e utilizam dois mecanismos diferentes para estabilizar esse intermediário tetraédrico. A aspartil (Asp) e metalo peptidases usam resíduos do sítio ativo para ativar moléculas de água para ataque nucleofílico, mecanismo conhecido como não covalente ácido-base, enquanto as cisteíno (Cys), serino (Ser) e treonino (Thr) peptidases usam resíduos do sítio ativo para hidrolisar as ligações peptídicas, nesse caso, o mecanismo envolve uma ligação covalente estável. Na catálise covalente, as histidinas funcionam normalmente como uma base, enquanto que na catálise não covalente, o resíduo de Asp e o zinco, no caso das metalo peptidases agem como ácidos e bases (TURK, et al., 2006). Esse estado de transição é um pré-requisito para a cisão da ligação peptídica em todas as classes com mecanismo conhecido (DRAG e SALVESEN, 2010).

De acordo com TANAKA (2002), têm sido estudadas maneiras de compreender melhor a seletividade e especificidade que as enzimas proteolíticas exibem, através de uma estrutura química ou proteica que apresente complementaridade entre o sítio ativo e o estado de transição e que possa catalisar a clivagem de ligações peptídicas nas mesmas condições da enzima natural. O intermediário tetraédrico pode ser esquematizado, conforme ilustra a Figura 1.6.

FIGURA 1.6 - Intermediários tetraédricos formados durante a clivagem peptídica



Fonte: DRAG e SALVESEN (2010)

As peptidases são encontradas em diferentes organismos sendo que em humanos existem cerca de 600 proteases e em camundongos mais de 80 podem ser encontradas, representando cerca de 2% dos genes expressos (TURK et al., 2012a).

Essas enzimas controlam uma variedade de processos fisiológicos essenciais para a vida como, ciclo celular, morte celular, cicatrização de feridas, digestão de alimentos e proteínas. Sua ação deve ser controlada, uma vez que os desequilíbrios em suas atividades ocasionam uma série de patologias, como doenças cardiovasculares, inflamação, câncer e doenças neurodegenerativas (TURK, 2006; DRAG e SALVESEN, 2010; TURK et al., 2012a).

1.3 Cisteíno peptidases

As catepsinas estão distribuídas entre as famílias das aspartil peptidases (catepsinas D e E), serino peptidases (catepsinas A e G) e as cisteíno peptidases (catepsinas B, C, F, H, K, L, O, S, V, W e X) e são expressas em diferentes células (TURK et al., 2012b).

De acordo com a classificação feita pelo MEROPS, um banco de dados contendo classificações hierárquicas de peptidases, as cisteíno peptidases pertencem à família C1, clã CA, que tem como principal representante a papaína,

extraída do látex de *Carica papaya*, com sua atividade proteolítica conhecida desde 1880 (RAWLINGS et al., 2012). As cisteíno peptidases utilizam um grupo tiol nucleofílico para fazer a quebra das ligações peptídicas e, conforme a sua especificidade estrutural, são classificadas em famílias tais como as catepsinas, cruzaina, calpaínas e caspases. Estas enzimas são amplamente distribuídas nos organismos, estão relacionadas com diversos processos fisiológicos atuando no *turnover* de proteínas, e como moléculas reguladoras e sinalizadoras (TURK et al., 2012b).

O desequilíbrio nos eventos de expressão das catepsinas, ativação do zimogênio, a sua inibição por proteínas, e a sua degradação levam ao envolvimento dessas enzimas no desenvolvimento de doenças como o câncer e artrite reumatoide, entre outras (TURK et al., 2012b).

As catepsinas foram inicialmente descritas na degradação de proteína no interior dos lisossomos. Por causa disso, todas as peptidases intracelulares foram chamadas de catepsinas. Hoje em dia, sabe-se que todas as cisteíno peptidases lisossomais são catepsinas, mas nem todas as catepsinas são lisossomais ou cisteíno peptidases (TURK e GUNCAR, 2003; SOSNOWSKI e TURK, 2016).

Em eucariotos as catepsinas L (E.C 3.4.22.15) são sintetizadas como pré-pró-enzimas (zimogênio) constituídas de peptídeo sinal N-terminal (10 a 20 aminoácidos), um pró-peptídeo (38 a 250 aminoácidos) e a enzima madura (220 a 260 aminoácidos) (WIEDERANDERS, 2003). O processamento da enzima madura requer a retirada do fragmento N-terminal, uma região pró-peptídica contendo um motivo ERFNIN, envolvido na inibição da enzima madura (KARRER et al., 1993). Essa região se sobrepõe ao sítio ativo bloqueando o acesso do substrato no sítio ativo, no entanto, ela é indispensável para promover o correto enovelamento da proteína, funcionando como uma chaperona intramolecular (YAMAMOTO et al., 1999). O processamento começa logo após a tradução, o peptídeo sinal é removido durante a translocação da proteína no retículo endoplasmático, e posteriormente a ativação da enzima ocorre no ambiente ácido dos lisossomos através da clivagem proteolítica do pró-peptídeo presente na região N-terminal (TURK et al., 2003), produzindo assim a enzima madura. *In vitro*, a ativação é feita por pH ácido ou pela ação de outra peptidase.

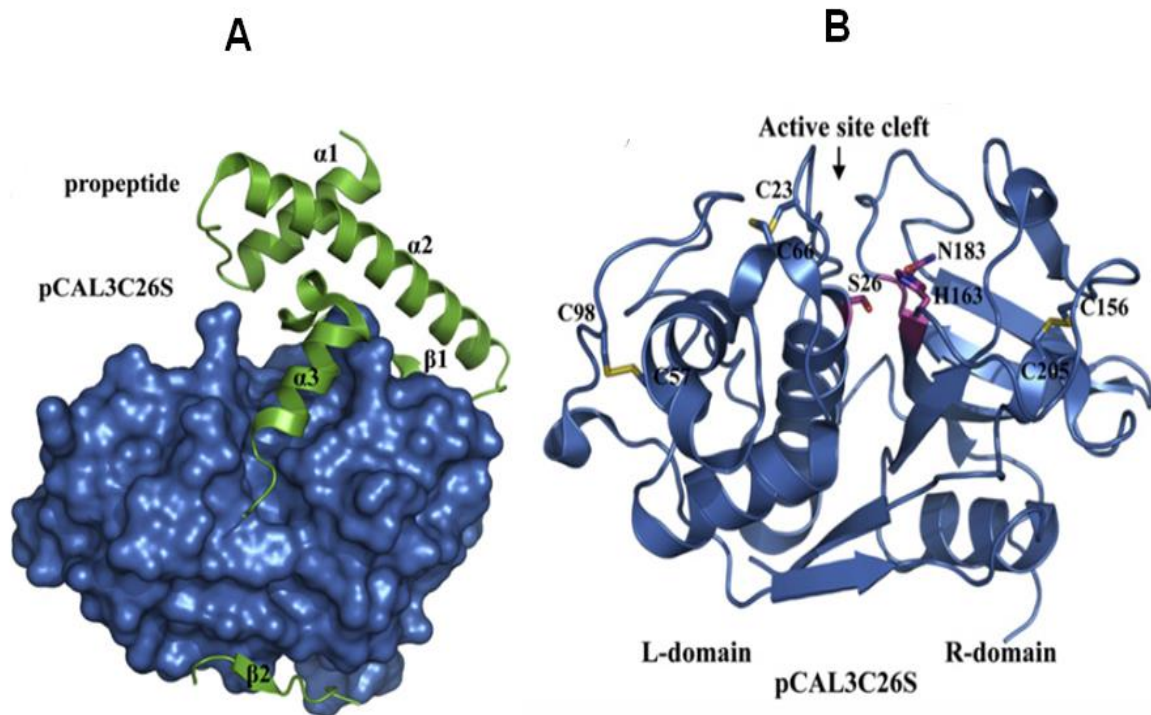
Essa subclasse das peptidases é caracterizada por ter uma conservada tríade catalítica Cys25, His159 e Asn175 e um resíduo de Gln19

(numeração da papaína) que auxilia na estabilização do intermediário oxianion da reação (BETON et al., 2012). O resíduo de cisteína presente no sítio ativo é característico das cisteíno peptidases com ação endopeptídicas (EC 3.4.22). Este resíduo promove a catálise nucleofílica, a qual leva a hidrólise da ligação peptídica no substrato (DAMASCENO et al., 2017).

Estruturas cristalográficas de catepsinas L têm sido elucidadas, (GUNCAR et al., 1999; WANG et al., 2006; STACK et al., 2008; HUANG et al., 2009; LARSON et al., 2009; BETON, et al., 2012). Todas apresentam enovelamento *papain-like*, são típicas representantes das cisteíno peptidases com ação endopeptídica, composta por dois lóbulos, o domínio L contendo α -hélices e o domínio R contando folhas β . Na interfase dos domínios, no topo da estrutura, há uma fenda (abertura) formando o sítio catalítico. O resíduo de cisteína da tríade catalítica se localiza no domínio L e o resíduo da histidina no domínio R, ambos acomodados na fenda catalítica. Estes dois resíduos catalíticos formam o par iônico tiolato-imidazólio, essencial para a atividade proteolítica das enzimas (TURK et al., 2012).

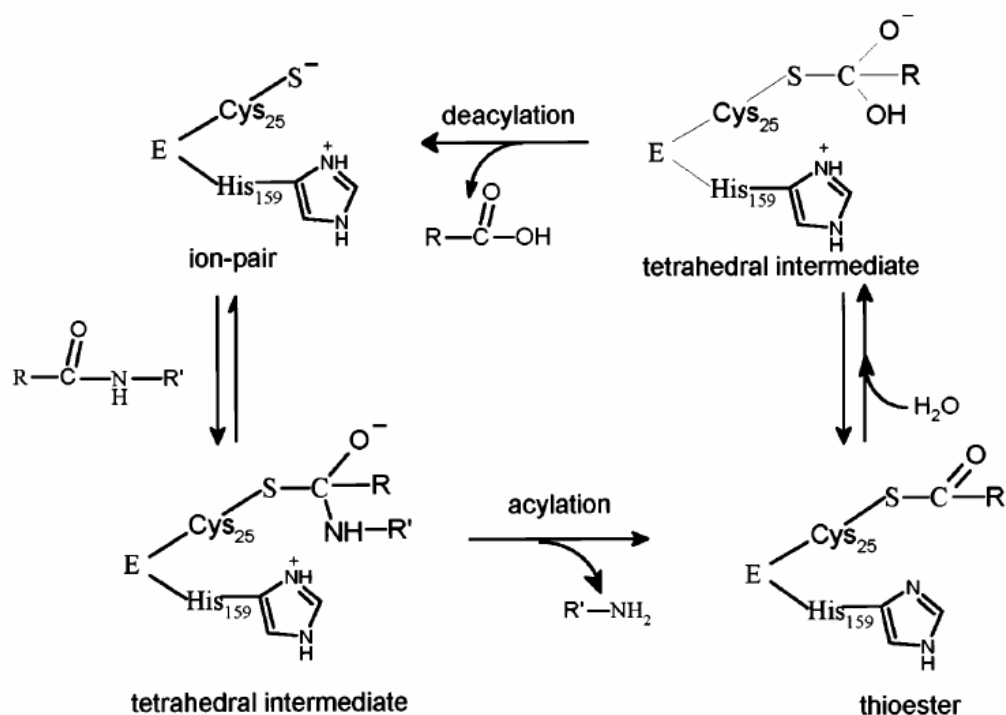
Na maioria das estruturas cristalográficas obtidas de catepsina L, a enzima está complexada com o pró-peptídeo. Em pH ácido a enzima sofre auto-ativação antes da cristalização. A recuperação da enzima ativa é baixa devido a auto degradação após a sua ativação, não sendo vantajoso retirar o pró-peptídeo da enzima madura para os ensaios de cristalização (LARSON et al., 2009). Estudos de mutagênese sítio-dirigida vêm sendo aplicados na obtenção das estruturas cristalográficas da catepsina L, onde sítios ativos mutantes são construídos mudando a cisteína catalítica pelo resíduo de serina para prevenir o processamento auto-catalítico (BETON et al., 2012). A Figura 1.7 ilustra uma estrutura cristalográfica de uma catepsina L digestiva do *Tenebrio molitor* obtida a partir da mutação no sítio catalítico. Em **A** – A enzima madura está representada por superfícies moleculares coloridas em azul e o pró-peptídeo colorido em verde. Em **B** - O resíduo catalítico mutante Cys26 \rightarrow Ser26 posicionado na região N-terminal no centro da α -hélice encontra-se no domínio L, enquanto os resíduos catalíticos His163 e Asn183 estão no domínio R. A cadeia lateral da cisteína apresenta as três ligações dissulfeto mostradas como bastões (amarelo). Essas ligações são encontradas nas enzimas da família C1, *papain-like*.

FIGURA 1.7 - Estrutura cristalográfica da pró-catepsina L e da catepsina L, enzima digestiva da larva do inseto *Tenebrio molitor*, denominada pCAL3C26S



Fonte: BETON, et al, (2012)

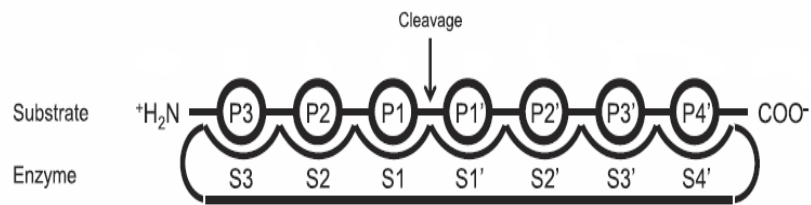
O mecanismo de ação das cisteíno peptidases na hidrólise dos peptídeos nos substratos pode ser visualizado na Figura 1.8 em um esquema simplificado. O grupo tiol da cisteína é desprotonado pela histidina adjacente. O enxofre aniônico se liga ao carbono da carbonila do substrato formando um intermediário tetraédrico, o qual é estabilizado pela cavidade oxianion sob a ação do resíduo de Gln, informação aqui não visualizada. Por um processo de acilação, este intermediário tetraédrico libera a porção C-terminal do substrato, deixando a enzima acilada (enzima-substrato tiol éster). Uma hidrólise forma um segundo intermediário tetraédrico e em seguida este intermediário sofre uma reação de desacilação, liberando a porção N-terminal e a enzima livre.

FIGURA 1.8 - Mecanismo da hidrólise do substrato pela enzima cisteína peptidase *papain-like*

Fonte: LECAILE et al., (2002)

Compreender as interações entre o sítio ativo da enzima e o substrato ainda são um desafio segundo TURK et al., (2012b). A papaína serviu como modelo no trabalho de SCHECHTER e BERGER (1967), quando propuseram uma nomenclatura para as posições dos resíduos no substrato P3, P2, P1 em direção ao N-terminal e P1', P2', P3' se direcionados ao C-terminal, e assim para os subsítios S3, S2, S1, S1', S2', S3', S4', locais de ligação da enzima com o substrato sobre a superfície da enzima. O P1 corresponde ao resíduo de aminoácido que antecede o ponto de clivagem no substrato, enquanto o P1' corresponde ao aminoácido no substrato, localizado após o ponto de clivagem, sugerindo que existiriam sete resíduos de substrato ligados em sete subsítios de S3 a S4'. Este modelo ainda é utilizado e ajuda a compreender a especificidade enzima-substrato. O modelo esquematizado pode ser visualizado na Figura 1.9. Variando os resíduos de aminoácidos nas posições P2, P1, P1' e P2' é possível descrever as especificidades dos subsítios S2, S1, S1' e S2', respectivamente.

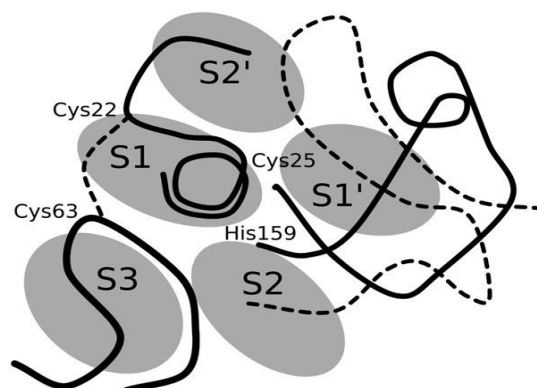
FIGURA 1.9 - Interação dos subsítios da enzima com os resíduos de aminoácidos do substrato



Fonte: Adaptado de DAMASCENO et al., (2017)

Três décadas depois, a definição dos locais de ligação do substrato foi revisada com base em uma visão fornecida por estruturas cristalinas dos complexos com moléculas pequenas (TURK et al., 1998). Essas estruturas mostraram que o substrato se liga ao longo da fenda ativa em uma conformação estendida, e os sítios de ligação ao substrato formam quatro segmentos na cadeia. Há a presença de duas alças curtas no domínio L e duas alças longas no domínio R. Uma terceira alça no domínio L pode ser chamada de dissulfeto (Cys22-Cys65), que conecta as duas alças ao longo da fenda do sítio ativo numa conformação estendida. Os resíduos do substrato P2, P1, P1' se ligam aos sítios de ligação dentro da fenda catalítica, assim a interação desses resíduos do substrato com a enzima envolvem tanto a cadeia principal como as cadeias laterais. Na Figura 1.10 observa-se à esquerda os subsítios S3, S1, S2' e à direita os S2 e S1' da fenda catalítica. De acordo com a numeração da papaína, os loops do domínio L incluem os Gln19-Cys25 e Arg59-Tyr67 enquanto os loops de domínio R contêm resíduos Leu134-His159 e Asn75-Ser205 respectivamente. Os resíduos do sítio ativo Cys25 e His159 e o dissulfeto Cys22-Cys63 estão indicados.

FIGURA 1.10 - Representação em 3D dos subsítios de ligação da enzima com o substrato ao longo da cavidade catalítica



Fonte: TURK et al., (2012b)

As catepsina L (EC 3.4.22.15) exibem uma preferência em hidrolisar ligações peptídicas com resíduos de aminoácidos hidrofóbicos, como o Phe em P2, sendo o subsítio S2 da cisteíno peptidase essencial para a tal especificidade. Enquanto que a catepsina B (EC 3.4.22.1) prefere Arg na mesma posição. Nesse contexto, o substrato fluorogênico Z-FR-MCA (carbобензоxi-Phe-Arg-7-amino-4-metil-coumarina) pode ser utilizado para distinguir uma catepsina da outra (TERRA e FERREIRA, 2005).

1.3.1 Inibidores das cisteíno peptidases

In vitro, a inibição da atividade proteolítica das enzimas da família *papain-like* pode ser feita irreversivelmente pelo peptídeo epóxido E-64. O mecanismo ocorre pela ocupação dos subsítios da enzima e alquilação da cisteína na tríade catalítica pelo grupo trans epóxido (SALVESEN e NAGASE, 2001). Visto que esse substrato reage com a cisteína do sítio reativo, ele é um inibidor bastante utilizado para a caracterização das cisteíno peptidases (TURK et al., 2012b).

O principal mecanismo de regulação das catepsinas maduras está na atuação dos inibidores proteicos endógenos como as cistatinas e as tiropinas, inibidores competitivos que se ligam fortemente, mas de maneira irreversível às cisteíno peptidases, impedindo a ligação do substrato ao sítio ativo (TURK et al., 2012b).

As cistatinas são uma superfamília de inibidores de cisteíno peptidase abrangendo um grande número de proteínas ativas contra a família *papain* peptidases. Desde a descoberta da primeira cistatina em ovo de galinha na década de 1960, o conhecimento sobre proteínas semelhantes presentes em mamíferos, aves, peixes, insetos, plantas e protozoários expandiu-se, e suas propriedades como potentes inibidores de peptidase foram firmemente estabelecidas. Atualmente, doze cistatinas em humanos já foram descritas e com base em detalhes estruturais distintos, mas também refletindo sua distribuição no corpo e funções fisiológicas, essa superfamília foi dividida em três famílias: as estefinas, as cistatinas e os cininogênios (ABRAHAMSON et al., 2003).

A família das estefinas não forma ligações de dissulfeto, possui massa molecular em torno de 11 kDa e atua no citosol. A segunda família é a das cistatina,

propriamente ditas, formada por proteínas que formam duas ligações de dissulfeto e têm massa molecular entre 13 e 24 kDa. A terceira família é a dos cininogênios compreendendo proteínas com elevada massa molecular entre 60 a 120 kDa, as quais são glicolisadas, formam ligações dissulfeto e estão envolvidas no processo de coagulação do sangue. (RAWLINGS e BARRETT 1990; TURK e BODE 1991; ABRAHAMSON et al., 2003). Os inibidores de cisteíno peptidases funcionam como uma proteção contra as peptidases lisossomais, liberadas ocasionalmente na morte celular, proliferação de células cancerosas ou por organismos invasores, como parasitas (DUBIN, 2004).

Em plantas, os inibidores de cisteíno peptidases são conhecidos como fitocistatinas (PhyCys), apresentam similaridade com a sequência de aminoácidos das cistatinas de animais, e se assemelham as estefinas por não formarem ligações dissulfeto. As fitocistatinas desempenham funções importantes na planta, que vai desde o controle de diversos processos fisiológicos e celulares à inibição de cisteíno peptidases exógenas secretadas por artrópodes, herbívoros e patógenos, que digerem e colonizam os tecidos das plantas (BENCHABANE et al., 2010).

Os inibidores que abrangem a família das fitocistatinas são identificados pela presença do motivo LARFAVDEHN formando uma α -hélice na região N-terminal (MARGIS et al., 1998). Em relação ao peso molecular, as fitocistatinas podem ser divididas em três grupos distintos: grupo I: proteínas de 12 a 16 kDa com apenas um domínio na sua estrutura proteica, a maioria das cistatinas de plantas estão inseridas nesse grupo; grupo II: proteínas com 23 kDa compartilhando um domínio N-terminal altamente conservado com as do grupo I e um segundo domínio C-terminal e grupo III: fitocistatinas de 80 kDa, conhecidas como multicistatinas por possuírem diversos domínios (MARGIS-PINHEIRO et al., 2008; WANG et al., 2008; CHU et al., 2011).

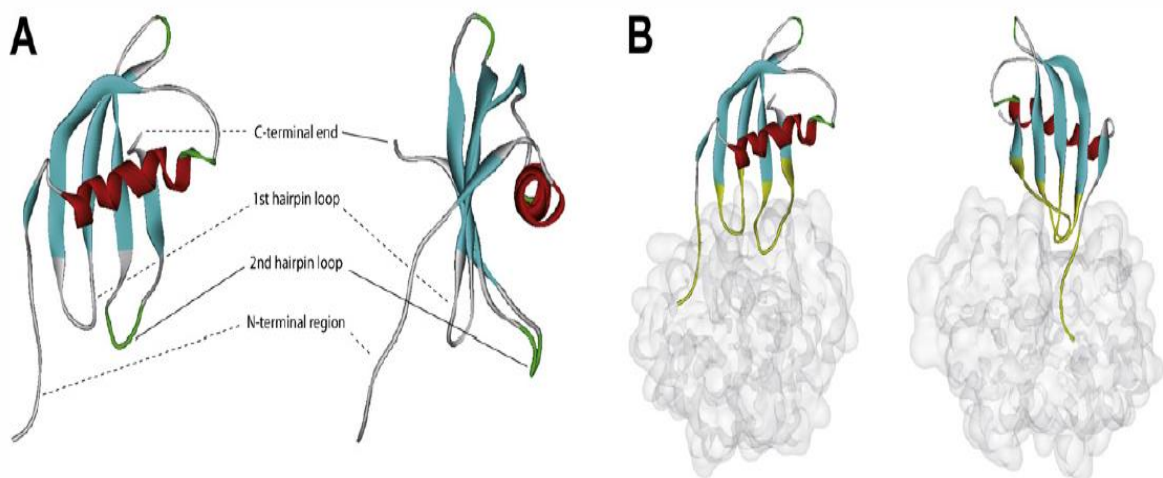
O domínio funcional das fitocistatina possui três motivos conservados para interação com cisteíno peptidase alvo: o motivo Q-X-V-X-G na região central da cadeia polipeptídica, onde Xaa é qualquer aminoácido, um dipeptídeo Pro-Trp ou Leu-Trp na região C-terminal, e um resíduo Gly conservado na região N-terminal (BARRETT et al., 1986 ; TURK e BODE 1999).

O inibidor de origem vegetal conhecido como orizacistatina isolada de arroz (*Oryza sativa*), pertencente ao grupo I das fitocistatina (ABE et al., 1987), mostrou ser eficiente na inibição da atividade proteolítica da papaína, e de outras

cisteíno peptidases (KONDO et al., 1990) , inclusive de cisteíno peptidases digestiva de insetos (LIANG et al., 1991).

Esse inibidor teve sua estrutura cristalográfica resolvida por NAGATA et al., (2000), sendo a primeira fitocistatina a ter estrutura resolvida. A estrutura tridimensional compreende 5 folhas- β antiparalelas encurvadas sobre uma α -hélice, a qual é estabilizada por um agrupamento hidrofóbico formado. As dobras das folhas- β formam as regiões de contato da cistatina com o sítio ativo da peptidase, formando um complexo equimolar fortemente ligado e reversível com a Cys peptidases, atuando como pseudo-substrato para adentrar na local ativo da enzima alvo, e bloqueando o acesso do substrato (BENCHABANE et al., 2010). Os três elementos estruturais do inibidor cistatínico que interagem com a cisteíno peptidase podem ser visualizados na Figura 1.11. Em **A** - é possível visualizar as partes lateral e frontal de orizacistatina que mostram um núcleo de folha de cinco cadeias (em azul), a α -hélice que abriga o motivo LARFAVDEHN conservado (em vermelho), a região N-terminal com o resíduo Gly conservado e os dois laços inibitórios abrigando os motivos conservados de Gln-Xaa-Val-Xaa-Gly (1st loop) e Pro-Trp (2nd loop), conhecidos por entrar na fenda do sítio ativo da enzima alvo. Em **B** – a vista lateral e traseira do modelo de encaixe para orizacistatina interagindo com o modelo Cys peptidase, *papain-like* (cinza claro).

FIGURA 1.11 – Modelo estrutural para a cistatina de planta, orizacistatina. Modelo construído por modelagem molecular



Fonte: BENCHABANE et al., (2010)

Com o intuito de potencializar o controle de insetos considerados pragas, os inibidores de peptidases presentes em plantas são objeto de estudos de diversos pesquisadores. Conhecida como agente de defesa da planta, essas proteínas, atuam contra as peptidases digestivas dos insetos podendo retardar o desenvolvimento larval. Essas observações levaram a investigações sobre o efeito da ingestão de cistatinas nas taxas de crescimento da população. Presume-se que a atividade antibiótica dos inibidores de peptidases deriva da interferência do inibidor com a digestão de proteínas, que diminui a disponibilidade de aminoácidos. A redução de aminoácidos poderia então reprimir diretamente a síntese de proteínas necessária para crescimento, desenvolvimento e reprodução (WOLFSON e MURDOCK, 1995).

Com a hipótese de que plantas usam inibidores de peptidases como mecanismo de defesa contra ataques de insetos e patógenos, foram realizados ensaios de inibição com cistatina purificadas de plantas transgênicas de cana-de-açúcar, os quais revelaram sua capacidade de inibir a atividade catalítica da cisteíno peptidase parcialmente purificadas do intestino médio do bicudo da cana-de-açúcar, um inseto que na sua fase larval se alimenta do rizoma da planta, levando à morte da mesma (RIBEIRO et al., 2008; SCHNEIDER et al., 2017).

As fitocistatinas de cana-de-açúcar foram descritas por REIS e MARGIS (2001), que identificaram, no banco de dados de ESTs (Expressed Sequence Tags) de cana-de-açúcar, o SUCEST (Sugarcane Expressed Tags Project), vinte e cinco fitocistatinas. Elas foram classificadas em quatro grupos: o grupo I compreende as cistatinas com organização estrutural típica das fitocistatinas, com as três regiões conservadas da superfamília cistatinas e a sequência consenso na região N terminal; o grupo II inclui as cistatinas com as mesmas características do grupo I, porém com uma extensão da região C terminal; o grupo III contém as que apresentam as regiões conservadas da superfamília das cistatinas, mas não possuem a sequência consenso característicos das fitocistatinas; e o grupo IV possui as que não apresentam nem a sequência consenso LARFAVDEHN na região N terminal e nem o motivo QXVXG conservado entre os membros da superfamília.

A primeira cistatina de cana-de-açúcar produzida de forma heteróloga foi denominada de canacistatina, CaneCPI-1, composta por 126 aminoácidos e massa molecular de 13 kDa (SOARES-COSTA et al., 2002), pertencente ao grupo I das canacistatinas. Hoje, seis canacistatinas recombinantes (CaneCPI-1, 2, 3, 4, 5 e

6) já foram caracterizadas e possuem validação de suas atividades inibitórias. A CaneCPI-1 apresentou atividade inibitória contra cisteíno peptidases, inclusive em catepsinas humanas (SOARES-COSTA et al., 2002). As CaneCPI-2, CaneCPI-3 e CaneCPI-4 mostraram ação inibitória sobre a papaína (GIANOTTI et al., 2005; GIANOTTI et al., 2008), a CaneCPI-4 também apresentou atividade inibitória contra catepsina B e L de humano (GIANOTTI et al., 2008) e as CaneCPI-1, 2, 3 e 4 apresentaram atividade inibitória contra a catepsina L digestiva, rSI-CathL do inseto praga da cana-de-açúcar, *Sphenophorus levis* (FONSECA et al., 2012). Recentemente foi validada a ação inibitória da CaneCPI-4 contra a cisteíno peptidase do protozoário *Plasmodium falciparum* (MELO et al., 2018).

2 OBJETIVOS

2.1 Objetivo Geral

Os objetivos deste trabalho foram a obtenção de uma catepsina L recombinante de *A. sexdens* (formiga saúva), a caracterização bioquímica e cinética e ensaios de inibição *in vitro* por canacistatinas.

2.2 Etapas para alcançar os objetivos

- 1) Amplificação e clonagem da ORF que codifica a síntese de catepsina de *A. sexdens* em vetor adequado para expressar a enzima;
- 2) Expressar a enzima em sistema de *E. coli*;
- 3) Purificar e caracterizar a enzima quanto à sua pureza, massa molecular, atividade enzimática e inibição utilizando os inibidores de cana-de-açúcar.

3 METODOLOGIA

3.1 Extração de RNA de *A. sexdens* e síntese de cDNA

A extração do RNA total foi feita pela Dra Ariele Cristina no Laboratório de Bioquímica e Biologia Molecular – UFSCar.

Para extração de RNA total, aproximadamente 120 mg de formigas (*A. sexdens*) foram maceradas em um procedimento utilizando TRIzol[®], procedimento este melhorado para a metodologia de extração de RNA em uma única etapa desenvolvido por CHOMCZYNSKI e SACCHI (1987). Seguiu-se o protocolo, sendo a única diferença a inclusão de mais uma etapa de purificação por partição com um volume de ácido fenol clorofórmio (mistura de fenol, clorofórmio, álcool isoamílico pH 4,3-4,7- Ambion). A amostra foi incubada por 1 minuto à temperatura ambiente e então foi centrifugada à 12000g à 4 °C por 15 minutos. Após esta etapa, a fase superior foi transferida para um tubo limpo e a metodologia foi seguida com a precipitação do RNA com isopropanol segundo o protocolo descrito pelo fabricante do TRIzol[®]. O *pellet* de RNA total foi ressuscitado em 40 µL de água tratada com DEPC (SAMBROOK e RUSSELL, 2001), e posteriormente foi incubada por 10 minutos a 60 °C.

A qualidade e quantificação de RNA total isolado foram estimadas pelas relações das absorvâncias da amostra medidas em 230, 260 e 280nm (A_{260}/A_{280} e A_{260}/A_{230}) utilizando BioSpec-nano (Shimadzu). A integridade da amostra de RNA foi verificada por gel de agarose 1 % em tampão TAE em água tratada com DEPC contendo brometo de etídio (0,0001 mg/mL). O gel foi submetido à eletroforese a 120 V por 30 minutos e posteriormente visualizado e fotografado pelo sistema de foto documentação computadorizado (Gel DocTM Ez System – Bio-Rad).

A partir do RNA total extraído foi obtido o cDNA (DNA complementar): em um tubo de PCR foram adicionados 4 µg de RNA total, 2 µL de Oligo(dT) *primer* 0.0005 mg/mL (Promega) e o volume foi completado para 13 µL com água tratada com DEPC. A amostra foi incubada por 5 minutos a 65 °C e resfriada imediatamente após o tempo de incubação. Então foi adicionada 4 µL de 5X *Reaction Buffer*, 2 µL de uma mistura de dNTPs 10 mM cada (Thermo Scientific) e 1 µL de RevertAid Reverse Transcriptase (Thermo Scientific). A amostra foi homogeneizada e incubada

por 60 minutos a 42 °C. A reação foi interrompida por aquecimento a 70 °C por 10 minutos.

3.2 Clonagem da ORF

Para obtenção da ORF de interesse foi necessária a construção por síntese química dos oligonucleotídeos de DNA complementares às extremidades de cada fita do cDNA. Estes oligonucleotídeos são os iniciadores da síntese de DNA *in vitro*, catalisada pela DNA polimerase. O desenho dos oligonucleotídeos foi realizado baseado na sequência codificadora de uma catepsina L da formiga cortadeira *Acromyrmex echinator*, disponível no banco de dados do *National Center for Biotechnology information* (<http://www.ncbi.nlm.nih.gov/>), sob o número de acesso EGI69825.1. A sequência de nucleotídeos da predita catepsina L de *A.echinator* está representada na Figura 3.1.

FIGURA 3.1- Sequência de nucleotídeos

```

ATGAAGAGTGTGGTGGTCTTGTTGTTACTGGCAGTTTTTACCACGGGACAAGCTATTTCCCTTCAACAAGA
TTTTGGATGCAGAATGGTTCATTTTTAAGACACATCATAAAAAAATCTATAAATCTTCGGTTGAGGAGGG
ATATAGAATGAAAATCTTCTTGGATAACAAGCGTAAAATTGCTGAACACAATCGCAAATATGAGCTGAAT
GAAGTCCCTACAAATGGGAATGAACAAATACGGTGACATGTTGCATCATGAATTTGTCAATACTTTGA
ATGGTTTTCAATAAGTCAGAAAAGGCTCAAAAGCAATTTATGGGAGCAACATTTATCTCGCCAGCTAATGT
CGAATTACCAAAAAGAAGTTGATTGGAGGAAACATGGTGCTGTTACAGAAGTTAAGGATCAAGGACATTTGT
GGCTCTTGTTGGGCATTTCTACTACTGGCTCTTTAGAAGGACAGCATTTCAGACAGACCGGTATTTTAG
TATCACTTAGTGAGCAAAATTTGATTGATTGTTCTGGTAACTATGGTAACGAAGGATGTAATGGTGGTTT
GATGGATAATGCCTTTAAATACGTCAGAGATAACAAAGGTCCTTGATACCGAAAAGAGTTACCCATATGAA
GCAGAGAATGACAAGTGCAGATAACAATCCGAGAAATAGTGGCGCTATTGATACCGGTTTTCGTAGATATTC
CAAGAGGAAATGAGCATAAGTTAAAGGCTGCAGTTGCCACTATAGGTCCGGTCTCGGTTGCTATTGATGC
CTCTCATGAATCTTTTCAGCTTTATAGTGAAGGAGTGTATTTTGATCCTGAGTGCGATTTCAGACAATTTA
GATCATGGTGTGTTGATCGTCGGCTATGGTACAGACTCAAAAACCGGTCATGATTATTGGCTGGGGAAGA
AAAGTTACATCAAAATGGCCAGAAATAAGGACAATCATTGTGGTATAGCAAGCAGTGCCAGCTATCCTCT
CGTTTAA

```

Fonte: Própria

3.2.1 Amplificação do DNA que codifica a síntese da catepsina L

O cDNA sintetizado a partir do RNA total extraído da formiga *A. sexdens* (operária) e o par de oligonucleotídeos foram utilizados na reação de polimerização em cadeia (PCR). Na amplificação da ORF de interesse foi utilizada a *Pfu* DNA Polymerase (Thermo Scientific), que possui a função de catalisar a polimerização de nucleotídeo na direção 5' → 3', e desempenha a atividade de

exonuclease (revisora), capacitando a polimerase a corrigir erros de incorporação de nucleotídeo. A reação foi realizada em um termociclador Eppendorf Mastercycler utilizando o programa descrito na Tabela 3.1.

TABELA 3.1 - Condições dos ciclos da reação de PCR para amplificação do gene

	CICLOS	T(°C)	TEMPO
Início	1X	95°C	3min
Desnaturação	38X	95°C	1min
Anelamento		55°C	1min
Extensão		72°C	4min
Finalização	1X	72°C	10min

Fonte: Própria

Para a amplificação da ORF de interesse foram utilizados 1 µl de cDNA, 0,2 mM de dNTPs, 1,0 µM de cada par de oligonucleotídeo (*sense* e *reverse*), 0,625 U de *Pfu* DNA Polymerase (Thermo Scientific) e 2,5 µl de *Pfu* Buffer 10X com MgSO₄ para uma reação de volume final de 25 µl. A reação de amplificação está descrita na Tabela 3.2.

TABELA 3.2 - Condições da reação de amplificação do DNA

Reagentes	Reação (µL)	Branco (µL)
cDNA	1,0	_____
<i>Primer forward</i> (10µM)	2,5	2,5
<i>Primer reverse</i> (10µM)	2,5	2,5
dNTP mix (10mM)	2,5	2,5
Tampão10XMgSO ₄ (25mM)	2,5	2,5
<i>Pfu</i> (2,5 U/µL)	0,25	0,25
H ₂ O milliQ	13,75	14,75
Volume total	25,0	25,0

Fonte: Própria

O produto da amplificação foi analisado aplicando-se 5 μL das amostras em gel de agarose 1% contendo $0,025\mu\text{g.ml}^{-1}$ de brometo de etídio em tampão TAE (80 mM Tris-HCl, 40 mM ácido acético, 2,5 mM EDTA, pH 8,3). Posteriormente, o volume total da reação foi aplicado em um gel de agarose, e o fragmento de DNA foi recortado e purificado do gel de agarose utilizando o kit Wizard SV gel and PCR clean-UP System (Promega) de acordo com o protocolo do fabricante. O produto purificado foi quantificado em espectrofotômetro BioSpec-nano (Shimadzu) (absorbância à 260 nm) e enviado para o sequenciamento. O sequenciamento foi feito pelo Centro de Pesquisas sobre o Genoma Humano e Células-Tronco Instituto de Biociências - USP -Setor de Sequenciamento de DNA. O centro de pesquisa utiliza o ABI 3730 DNA *Analyser*, um sistema de análise de DNA de 48 capilares com a tecnologia Life Technologies – Applied Biosystems. O resultado do sequenciamento foi analisado por inspeção visual dos eletroferogramas, e em seguida foi feito um alinhamento no banco de dados do NCBI utilizando o Blastx com a função de buscar identidade com sequências das formigas depositadas no banco de dados.

3.2.2 Análise qualitativa da expressão da ORF

Com o objetivo de saber se o gene que codifica a ORF da catepsina na formiga cortadeira *A.sexdens* seria expresso nos diferentes estágios de vida do inseto (larva, pupa, adulto) assim como em algumas partes da formiga adulta conhecida como operária (cabeça, tórax, gaster), foram feitas PCR tendo como molde o cDNA sintetizado a partir do RNA total extraído dessas várias amostras. Os oligonucleotídeos *forward* e *reverse* sintetizados, conforme descrito anteriormente, foram utilizados na reação de amplificação. As condições dos ciclos da reação da PCR, assim como as condições da reação de amplificação do DNA, estão descritas no item 3.2.1.

3.2.3 Extração do DNA genômico

Devido à relação simbiótica entre a formiga *A.sexdens* e o fungo *L. gongylophorus*, surgiu a necessidade de saber se o gene que codifica a catepsina

era do inseto ou se era do fungo, uma vez que a formiga se alimenta desse micro-organismo. Com esta finalidade, foi feita a extração do DNA genômico da cabeça da formiga, parte do corpo que poderia ter menos contato com o fungo e assim menor probabilidade de contaminação pelo mesmo.

Cerca de 0,06 mg do material foi lavado cuidadosamente com etanol 70% e em seguida com H₂O *milliQ* autoclavada por 3 vezes, para remoção de resquícios de fungos e bactérias. Em seguida a amostra foi macerada com ajuda de um bastão de vidro para uma melhor homogeneização. A amostra foi ressuspendida em 462,5 uL de tampão T.E. (Tris-EDTA, pH 8.0), 30 uL de SDS 10 % e para otimizar a lise foi adicionado 7,5 uL de proteinase K (20 mg/mL). A amostra ficou encubada por 1 hora no Eppendorf ThermoMixer[®] (banho seco) a 60 °C para a ação da proteinase K. Em seguida, foi acrescentado o mesmo volume da reação 500 uL de fenol/clorofórmio (preparado na proporção de 1:1 em volume), a amostra foi homogeneizada por inversão e centrifugada a 14000 g por 5 minutos a 4 °C.

Após essa etapa, a parte superior (fração aquosa, que contém o DNA) foi transferida para outro tubo de 2 mL. Nesse momento, tomou-se o cuidado para transferir apenas a parte superior, não coletando nem a interfase, que contém proteínas precipitadas e, tampouco a fração inferior, que possui compostos orgânicos dissolvidos no fenol/clorofórmio. A etapa de extração foi repetida 3 vezes, com o propósito de obter uma amostra mais pura, ausente de contaminantes como proteínas. Para remoção dos restos de fenol da amostra, acrescentou-se à reação o mesmo volume de clorofórmio. A amostra foi homogeneizada, centrifugada a 14000 g por 5 minutos a 4 °C e a fase aquosa foi transferida para um tubo de 1,5 mL.

O DNA extraído foi precipitado com 0,1V de Acetato de sódio 3 M e 0.6V de Isopropanol. A amostra foi homogeneizada até a precipitação do DNA e, logo após, centrifugada a 5000 g por 1 minuto e o sobrenadante foi removido. Em seguida o pellet foi lavado com etanol 70 % gelado, centrifugado novamente a 5000 g por 1 minuto e o sobrenadante foi removido.

O DNA foi seco à temperatura ambiente por um período de 1 hora, tomando-se o cuidado para a amostra não desidratar. Para finalizar, após o período de secagem, foi adicionado ao tubo de 1.5 mL, contendo o DNA extraído, 50 µl de T.E. (100 mM Tris, 200 mM EDTA, pH 8.0). Um gel de agarose 1% foi feito para análise do DNA genômico, e a quantificação foi feita em um espectrofotômetro BioSpec-nano (Shimadzu), absorvância a 260 nm.

Após extração e confirmação em gel de agarose, o DNA genômico serviu como molde na amplificação por PCR, utilizando os oligonucleotídeos *forward* e *reverse* sintetizados e descritos no item 3.1. As condições dos ciclos da reação da PCR, assim como as condições da reação de amplificação do DNA estão descritas no item 3.2.1, alterando apenas a quantidade de DNA adicionado, nesse caso cerca de 2 µl de DNA genômico (430 ng/µl).

3.2.4 Digestão do vetor pET32a-c(+) e do produto de PCR

Para a expressão heteróloga de cisteína peptidase da família *papain-like*, a bactéria *E.coli* geralmente é a primeira escolha devido ao seu manuseio simples, meio de cultura barato e alto rendimento. Estas vantagens compensam o esforço adicional quando o produto precipita em corpos de inclusão. As principais preocupações para a produção de peptidase em *E.coli* são assegurar taxas elevadas de transcrição e tradução, controlar a toxicidade do produto gênico heterólogo, manter a solubilidade da peptidase e prevenir a degradação proteolítica da proteína expressa. A maioria dessas preocupações podem ser abordadas pela seleção de vetores de expressão com promotores fortemente indutíveis e fortes sítios de ligação ribossômica, sequências de direcionamento que direcionam o produto gênico para o compartimento periplasmático, e a introdução de peptídeos *tags* que permitem fácil purificação (BROOME et al., 2004; KRAMER et al., 2007).

O vetor de expressão pET32a-c(+) escolhido para clonagem, permite a expressão de altos níveis de sequências peptídicas fusionadas com 109 aminoácidos da Tiorredoxina (Trx•Tag™), proteína que auxilia na redução da formação de proteínas agregadas, aumentando a solubilidade da proteína expressa. O vetor dispõe de sítios de clonagem que permitem produzir proteínas em fusão com as sequências His-Tag® e S-Tag™ para detecção e purificação e carrega o gene AmpR que confere resistência ao antibiótico ampicilina. Na Figura 3.2 em destaque as enzimas de restrição utilizadas na clonagem (em vermelho).

Figura 3.2 - Sequência do sítio de clonagem múltipla dos vetores pET-32a-c(+)



Fonte: Adaptado de <https://www.snapgene.com/>

O vetor e ORF foram digeridos separadamente com as enzimas de restrição *EcoRI* (Thermo Scientific) e *NcoI* (Thermo Scientific), conforme sugerido pelo fabricante. As reações de digestão foram mantidas a uma temperatura de 37 °C por 4 horas. Os produtos das reações de digestão foram analisados em gel de agarose 1% e purificados utilizando o Kit Wizard SV Gel and PCR Clean-UP System (Promega) como descrito pelo fabricante. As condições da reação de digestão estão descritas na Tabela 3.3. Tanto o plasmídeo como o inserto apresentou tamanho em pares de bases esperado. O produto da reação de ligação inserto-pET32a foi transformado em linhagem *DH5α* de *E.coli* e colônias obtidas foram inoculadas em meio LB (*Lysogeny broth*) para posterior extração do DNA plasmidial.

TABELA 3.3- Condições da reação de digestão *EcoRI* e *NcoI*

	pET-32 a (+)	Produto de PCR
<i>EcoRI</i> 10 U/ µL	1,0 µL	1,0 µL
<i>NcoI</i> 10 U/ µL	1,0 µL	1,0 µL
<i>Green buffer</i> 10X	2,0 µL	2,0 µL
DNA	6,5 µL (1,0 µg)	7,0 µL (0,5 µg)
H ₂ O	9,50 µL	9,0 µL
Volume total	20,0 µL	20,0 µL

Fonte: Própria

Posteriormente, foi feita a reação de ligação do vetor com o inserto (produto de PCR). A reação foi montada de acordo com a proporção vetor e inserto 1:3, com um volume final de 20,0 μL . A reação foi feita na presença da T4 DNA *Ligase* (Invitrogen) e foi incubada a 4 °C por 16 horas. As condições da reação de ligação estão descritas na Tabela 3.4.

TABELA 3.4 - Condições da reação de ligação

Reação	Ligação (1:3)*
pET-32 a (+) (22,70 ng/ μL)	4,40 μL
Produto PCR (13,34 ng/ μL)	3,70 μL
T4 DNA <i>Ligase</i> (5 U/ μL).	0,2 μL
10X T4 DNA <i>ligase buffer</i>	2,0 μL
H ₂ O	9,70 μL
Volume total	20,0 μL

*Proporção vetor / inserto

Fonte: Própria

O produto de ligação, denominado de AsCathL, foi utilizado para transformar células competentes de *E.coli* da linhagem *DH5 α* . O processo de transformação utilizado foi por choque térmico, seguindo o protocolo modificado desenvolvido por HANAHAN (1983). Um volume de 10 μL do produto da ligação foi adicionado às células competentes e a mistura foi mantida no gelo por 30 minutos. Após a incubação, foi realizado o choque térmico por 2 minutos a 42 °C. Em seguida o tubo (Eppendorf) de 1,5 mL contendo a reação foi mantido por 2 minutos no gelo, após esse período foram adicionados 400 μL de meio LB (1% (m/v) triptona, 0,5% (m/v) extrato de levedura 1% (m/v) e cloreto de sódio (pH 7,5) à reação e, em seguida incubada, a 37 °C por 60 minutos sob agitação de 250 rpm.

Após o tempo de incubação, a mistura foi centrifugada por 30 segundos e cerca de 400 μL do sobrenadante foi desprezado. O produto da transformação, aproximadamente 100 μL , foi plaqueado em 2 placas de Petri individuais 25 μL e 75 μL respectivamente. As placas continham 25 mL de meio LB com 1,5 % (m/v de ágar) e 100 $\mu\text{g/mL}$ de ampicilina. Após o plaqueamento, as placas foram incubadas em estufa a 37° C por 16 horas. As colônias transformantes foram isoladas e, em seguida, inoculadas para extração do DNA plasmidial.

Para analisar o plasmídeo contendo a ORF de interesse, colônias obtidas na etapa anterior foram selecionadas e cultivadas em meio LB, contendo 0,1

mg/mL de ampicilina sob agitação de 250 rpm a 37 °C por 16 horas. A extração do DNA plasmidial foi feita com a utilização do kit Wizard® Plus SV Miniprep (Promega) de acordo com o protocolo sugerido pelo fabricante. Um gel de agarose 1

% foi realizado para análise do DNA plasmidial, e a quantificação do DNA foi feita em um espectrofotômetro BioSpec-nano (Shimadzu), absorvância a 260 nm. A confirmação do clone positivo foi feita por PCR, tendo como DNA molde o DNA plasmidial extraído e o uso dos mesmos oligonucleotídeos empregados na amplificação do inserto. Os reagentes da reação de amplificação do clone estão descritos na Tabela 3.5.

Tabela 3.5 - Condições da reação de amplificação do clone.

Reagentes da	Reação (µL)	Branco (µL)
DNA plasmidial (108 ng/ µL)	1,0	_____
<i>Primer foward</i> (10 µM)	2,5	2,5
<i>Primer reverse</i> (10 µM)	2,5	2,5
dNTP mix (10 mM)	2,5	2,5
Tampão 10X com (NH ₄) ₂ SO ₄)	2,5	2,5
<i>Taq DNA Polymerase</i> (2,5 U/µL)	0,25	0,25
<i>MgCl₂</i>	2,0	2,0
<i>H₂O milliQ</i>	11,75	12,75
Volume Total	25,0	25,0

Fonte: Própria

O clone positivo foi analisado por restrição, utilizando as enzimas *NcoI* e *EcoRI* sob as condições descritas na Tabela 3.3. Para confirmação da clonagem, a os *primers T7 promoter* e *T7 terminator* necessários para o sequenciamento da fita simples. A amostra, juntamente com os *primers*, foi enviada para análise sequencial. O sequenciamento foi feito pelo Centro de Pesquisas sobre o Genoma Humano e Células-Tronco Instituto de Biociências - USP -Setor de Sequenciamento de DNA. O resultado do sequenciamento foi analisado por inspeção visual dos eletroforegramas, e em seguida feito um alinhamento no banco de dados do NCBI utilizando o Blastx. A predição da estrutura primária da enzima foi realizada a partir da tradução da sequência de nucleotídeo utilizando o software online (<http://web.expasy.org/translate/>).

3.3 Expressão recombinante da AsCathL em *E.coli ArcticExpress*

Após o sequenciamento e confirmação da inserção da ORF, a linhagem de *E.coli ArcticExpress* (DE3)RIL foi transformada com a construção AsCathL, usando o método do choque térmico com adição do β - Mercaptoethanol, procedimento realizado de acordo com o protocolo do fabricante.

Cerca de 50 ng (0,5 μ L) do DNA plasmidial contendo o gene de interesse foi misturado cuidadosamente em 100 μ L de célula competente *ArcticExpress* (DE3)RIL e, em seguida, a reação foi incubada no gelo por 30 minutos. Após a incubação, a reação foi submetida a um banho a 42 °C por 20 segundos e novamente a reação foi incubada no gelo por mais 2 minutos. Posteriormente, adicionou-se cerca de 900 μ L de meio SOC (meio LB + glicose), pré-aquecido a 42 °C na reação de transformação e esta foi incubada a 37 °C por 1 hora sob agitação de 250 rpm.

Após o período de incubação, as células transformadas, foram concentradas por centrifugação a 200 g por 4 minutos a 4 °C e, em seguida, ressuspensas em 200 μ L de meio SOC. 100 μ L das células foram inoculadas em placas Petri contendo 25 mL de meio LB ágar, 100 μ g/mL de ampicilina e 20 μ g/mL de gentamicina, antibióticos necessários à seleção das células de *ArcticExpress*. A placa foi incubada em estufa a 37 °C por 16 horas. Para controle da reação transformante, foi utilizado 1 μ L do plasmídeo pUC18 em 100 μ L de célula competente.

As colônias foram selecionadas e inoculadas individualmente em 5 mL de meio LB contendo os antibióticos anteriormente citados. As culturas foram incubadas a 37 °C por 16 horas sob agitação constante (250 rpm). No dia seguinte, foi retirada uma alíquota de 400 μ L de cada meio de cultura e a cada uma adicionou-se 400 μ L de glicerol 50% para armazenamento a -80 °C. O restante do meio de cultura foi utilizado para obtenção do DNA plasmidial utilizando o kit Wizard® Plus SV Miniprep (Promega). Um gel de agarose 1% foi realizado para análise e em seguida o DNA plasmidial foi quantificado em um espectrofotômetro BioSpec-nano (Shimadzu), absorvância a 260 nm. Posteriormente uma PCR, sob as condições citadas na Tabela 3.5, foi feita para verificação das colônias transformantes.

Após confirmação por PCR, a alíquota estocada a -80 °C, foi estriada em LB ágar com os antibióticos apropriados (100 μ g/mL de ampicilina e 20 μ g/mL de

gentamicina). A placa foi incubada em estufa a 37 °C por 16 horas. Posteriormente, uma colônia isolada foi inoculada em 5 mL de meio LB contendo os antibióticos necessários. A cultura foi mantida a 37 °C sob agitação constante de 250 rpm por 16 horas. Desse meio de cultura, foi pipetado 100 µL em um tubo Falcon de 50 mL contendo 5 mL de meio LB sem adição dos antibióticos.

A cultura permaneceu a 30 °C sob agitação constante de 250 rpm por 3 horas, até que a densidade óptica DO_{600} atingisse um valor entre 0.4 – 0.6. Após as 3 horas, 100 µL do meio de cultura foi recolhido para posterior análise em gel de SDS-PAGE. O meio de cultura foi incubado a 10 °C por 10 minutos a 250 rpm para ambientalização da cultura bacteriana nessa temperatura e, logo após, foi adicionado o indutor IPTG (Isopropil-β-D-tiogalactopiranosídeo) com concentração final de 0.2 mM. O meio de cultura foi incubado por 24 horas a 10 °C sob agitação de 250 rpm. Após o período de incubação, foi retirada uma alíquota de 100 µL do meio de cultura para posterior análise em gel SDS-PAGE para controle da amostra induzida. O meio de cultura foi centrifugado a 6000 x g por 10 minutos a 4 °C para separação da célula do sobrenadante.

As células foram ressuspensas em 5,0 mL de tampão de lise (50 mM Tris-HCl pH 8,0). A lise mecânica das células foi feita pelo ultra-som (Sonic Dismembrator Fisher Scientific mode 500) com 23 % de amplitude durante 2 minutos, sendo 5 segundos de sonicação e 10 segundos de repouso. Posteriormente, o lisado foi centrifugado a 4 °C por 20 minutos a 10000 x g para separação da fração solúvel e insolúvel. Alíquotas do precipitado celular obtido nesta etapa e do sobrenadante foram retiradas para análise em gel desnaturante de poliacrilamida SDS-PAGE 15 % segundo LAEMMLI (1970).

3.4 Análise da expressão da pró-enzima por *Western blot*

Para detectar a pró-enzima expressa fusionada com a His-Tag, foi feito um *immunoblotting* utilizando anticorpo anti-His. No procedimento do *western blot* as amostras do meio de cultura foram aplicadas em gel desnaturante de poliacrilamida. As proteínas separadas no gel foram transferidas para uma membrana de nitrocelulose. Para tanto, montou-se um “sanduíche” colocando-se o gel em contato com a membrana de nitrocelulose, e de cada um dos lados, um pedaço de esponja, todos imersos em tampão de transferência (200 mM Tris, Glicina 50 mM, Metanol

15%). O “sanduíche” foi colocado no Trans-Blot Turbo (BIO-RAD) e submetido a uma voltagem de 25 V e 1.0 A, durante 30 minutos. Após a transferência, a membrana foi corada com solução Ponceau (0,5 % Ponceau em ácido acético 0,1 %) por 5 minutos sob agitação, e lavada, em seguida foi lavada com água destilada até que fosse possível visualizar as bandas. A membrana foi então incubada por 1h, à temperatura ambiente, em solução de bloqueio (tampão TBS 50 mM Tris, 150 mM NaCl, pH 8.0), acrescido de leite em pó desnatado 5 %. Após esse intervalo, a membrana foi lavada 3 vezes com TBS, acrescido de *Tween* 20 0,3 % (5 minutos cada lavagem). A membrana foi incubada por 1 hora com anticorpo anti-His (diluído na proporção 1:1000 em TBS). Após a reação com o anticorpo a membrana foi colocada em uma solução reveladora (RENART, et al., 1979).

3.5 Purificação da pró-enzima

A fração solúvel foi purificada por cromatografia de afinidade em metal por fluxo de gravidade, usando uma coluna contendo 1 mL de resina de níquel NI-NTA *Superflow* (Qiagen).

Antes da etapa de purificação, a resina foi lavada com água *milliQ* e equilibrada com 5 CV (5mL) de tampão A (50 mM de Tris-HCl, pH 8,0). Em seguida, 20 mL da fração solúvel foram aplicados na coluna, seguido por uma etapa de lavagem com 30 CV de tampão A acrescido de 50 mM de NaCl. Após a lavagem, a eluição das proteínas foi realizada com 1 mL de tampão A acrescido de 250 mM de Imidazol. Este reagente possui forte afinidade pelo níquel, competindo com a cauda de histidina presente na proteína recombinante, permitindo que a proteína se solte do metal e então seja eluída. Após essa etapa, foi necessário fazer a troca de tampão para a retirada do Imidazol da solução, para que não houvesse interferência do sal no ensaio enzimático. A troca de tampão foi feita usando um sistema de ultracentrifugação (Amicon Ultra-15 Centrifugal Filter Devices), com molecular *weight cut off* (MWCO) de 10 kDa. Após cada 5 minutos sob uma rotação de 5000 x g, 1 mL de tampão A era adicionado a fração eluída, processo repetido 4 vezes. Posteriormente, a fração dessalinizada foi concentrada 2 vezes utilizando o mesmo sistema de ultracentrifugação. A amostra foi submetida ao ensaio de atividade enzimática.

Uma segunda etapa de purificação foi adicionada ao processo. Assim, a amostra eluída da coluna de níquel, dessalinizada e com atividade enzimática foi submetida à segunda etapa de purificação fazendo uso da cromatografia de troca iônica, utilizando uma coluna de 1 mL *Hitrap* QFF acoplada ao sistema FPLC (*Fast protein Liquid Chromatography*), utilizando o equipamento AKTA (GE Healthcare Sciences). A coluna foi previamente equilibrada com 5 mL de tampão A e o volume de 1 mL da amostra a ser purificada foi injetado na coluna contra o tampão A. O fluxo da aplicação e eluição da amostra foi de 1 mL min⁻¹.

A eluição da proteína foi feita em tampão A acrescido de 1 M de NaCl, em 30 CV. As amostras eluídas da coluna foram coletadas em frações de 1 mL e estas foram submetidas a um ensaio de atividade enzimática. A fração que apresentou atividade foi concentrada 4 vezes utilizando o sistema de ultracentrifugação mencionado anteriormente. Todas as etapas envolvidas no processo da purificação da pró-enzima foram analisadas por eletroforese em gel desnaturante de poliacrilamida SDS-PAGE 15 %, LAEMMLI (1970) e a quantificação feita pelo método de BRADFORD (1976).

3.6 Recuperação da pró-enzima a partir de corpos de inclusão

A expressão recombinante da proteína AsCathL realizada em *E.coli* resultou numa pequena fração solúvel, porém proporcionou uma expressão acentuada de proteína em corpos de inclusão, o que nos levou a “resgatar” essas proteínas através do *refolding*. A literatura disponibiliza vários protocolos, contudo, a estratégia geral é baseada em três passos: isolamento e lavagem dos corpos de inclusão, solubilização das proteínas agregadas e *refolding*. Nesse contexto, partimos em busca de uma metodologia que proporcionasse proteína solúvel e com atividade enzimática.

O *pellet* obtido após a lise descrito no item 3.2, foi lavado 2 vezes com tampão (20mM Tris HCl acrescido de 2.5 M de ureia, 2.5 % de Triton X-100, pH 8.0), seguido de centrifugação a 14000 g por 30 minutos a 4 °C. Após essa etapa, o pellet foi pesado para o cálculo de volume de tampão solubilizante a ser utilizado. Foi calculado o volume de cerca de 5 mL de tampão para cada 1g de corpos de inclusão. As proteínas agregadas foram solubilizadas em tampão (20 mM de Tris HCl, 6 M de Cloreto de guanidínio, 250 mM de NaCl, 20mM de Imidazol, pH 8.0),

sob suave agitação por 60 minutos a 4 °C, e centrifugadas a 14000 x g por 30 minutos a 4 °C.

Para a purificação da proteína recombinante, utilizou-se o método da cromatografia de afinidade por fluxo de gravidade Ni²⁺NTA. A resina foi lavada com água *milliQ* e equilibrada com 5 CV do tampão de solubilização. O sobrenadante foi incubado na coluna contendo a resina de níquel por 60 minutos à temperatura ambiente. Após esse tempo, a resina foi lavada com 10 CV de cada um dos seguintes tampões: Tampão 1 (20 mM de Tris HCl, 6 M de Cloreto de Guanidínio, 250 mM de NaCl, pH 8.0) ; Tampão 2 (20 mM de Tris HCl, 8 M de Ureia, 500 mM de NaCl, pH 8.0) ; Tampão 3 (20 mM de Tris HCl, 8 M de Ureia, 30 mM de Imidazol, pH 8.0). As proteínas ligadas à resina foram eluídas com tampão (20 mM de Tris HCl, 8 M de ureia, 1 M de Imidazol, pH 8.0), e quantificadas pelo método de BRADFORD (1976).

O método da diluição da proteína foi escolhido para conseguir a pró-enzima reenovelada. Foram adicionados 50 µg de proteína em solução em 50 mL de tampão de *refolding* (100 mM de Tris HCl, 20 % de Glicerol, 250 mM de L Arginina e 1 mM de Glutathione Oxidada, pH 8.0). A amostra ficou encubada por 16 horas em repouso a 4 °C. No dia seguinte, a amostra diluída foi concentrada até o volume de 2 mL em membrana de *cut-off* de 10 kDa (Amicon Ultra-15 Centrifugal Filter Devices), seguida de uma dessalinização (troca de tampão de para o tampão 50 mM Tris HCl, pH 8.0), utilizando o mesmo sistema de ultracentrifugação. Este protocolo de recuperação de proteína a partir de corpos de inclusão foi feito com base no método trabalhado por (SHENAI et al., 2000), com algumas modificações.

3.7 Ensaio enzimático para determinação da atividade proteolítica da cisteíno peptidase

3.7.1 Ativação da cisteíno peptidase

Para gerar uma cisteíno peptidase ativa (proteína madura), a pró-região com o domínio inibitório precisa ser removida, uma vez que a maioria das peptidases são expressas na forma de zimogênios (inativas) estáveis em pH alcalinos ou ligeiramente ácidos. A remoção da pró-região acontece por autocatálise

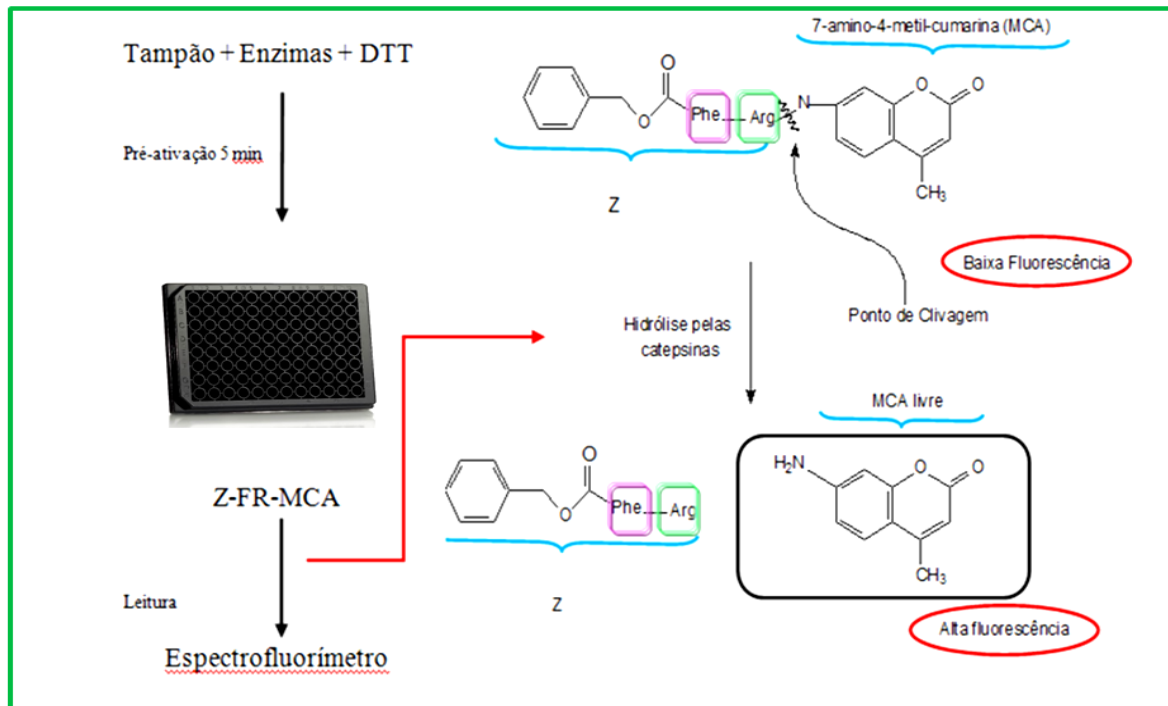
incubando-se o precursor da proteína madura em tampão ácido. Nesse trabalho a ativação foi feita segundo a metodologia empregada por BROMME et al., (2004), com algumas modificações.

Nesse contexto, os ensaios enzimáticos foram feitos incubando por 5 minutos 8 nM de enzima purificada, 5 μ M de DTT (Dithiothreitol) para redução do resíduo da cisteína do sítio catalítico, em tampão 100 mM Acetato de sódio, pH 4,5 numa reação com volume final de 200 μ l. O tampão acetato é comumente empregado nos ensaios de atividade proteolítica para catepsinas, e o pH ácido é responsável por enfraquecer as ligações entre a região pró-peptídica e o sítio ativo. O tempo de incubação é importante e deve ser o necessário para determinar o nível de atividade auto-catalítica, uma vez que a auto-catálise, assim como a ativação, leva a uma parcial degradação da enzima recombinante (BROMME et al., 2004).

A atividade enzimática da cisteíno peptidase foi determinada utilizando o substrato fluorogênico Z-FR-MCA (carbobenzoxi-Phe-Arg-7-amida-4-metilcumarina), assim, após os 5 minutos de incubação o substrato foi adicionado à reação na concentração final de 0,11 mM. Esse substrato é composto por peptídeo ligado à sonda fluorescente 7-amino-4-metil-cumarina (MCA), e quando hidrolisado pela enzima, o grupo fluorescente é liberado em solução permitindo a detecção da formação do produto. Este substrato é muito utilizado para o monitoramento da atividade de cisteíno peptidase (catepsina L) com atividade endopeptídica.

A leitura da atividade enzimática foi medida pelo aumento da fluorescência, ocasionada pela hidrólise do substrato Z-FR-MCA, monitorada continuamente por espectrofluorímetro (SpectraMAX[®] M3). Os ensaios enzimáticos foram feitos em triplicatas, utilizando placa de Elisa 96 poços, coloração preta, fundo plano, mantida em compartimento termostaticado a 23 °C, sendo a fluorescência detectada com fenda de excitação ajustada para λ_{ex} 380 nm e emissão para λ_{em} 460 nm. A metodologia geral está esquematizada na Figura 3.4. Destaque para o ponto de clivagem no substrato pela catepsina e a fluorescência da 7 amino 4 metilcumarina (MCA) livre.

FIGURA 3.3 - Esquema geral dos ensaios enzimáticos na presença do substrato fluorogênico Z-FR-MCA



Fonte: Adaptado de SEVERINO, 2008.

3.7.2 Caracterização cinética

A atividade catalítica da cisteína protease em função do pH foi analisada na presença de tampão 100mM de acetato de sódio com pH no intervalo de 3,5 a 5,5. Para a determinação da constante de Michaelis-Menten, K_m , foi feita uma diluição do substrato fluorogênico Z-FR-MCA, variando a concentração: 0,1; 1,9; 3,9; 7,8; 15,6; 31,25; 62,5; 125 e 250 μM . A constante de Michaelis-Menten (K_m) e os valores de velocidade máxima da reação (V_{max}) foram calculados pelo método da regressão não linear utilizando o software GraphPad Prism 5.0.

3.8 Ensaios de inibição da atividade proteolítica

No processo inicial da inibição proteolítica foi feita uma triagem para identificar possíveis inibidores capazes de inibir a atividade da catepsina L. Os inibidores analisados foram seis cistatinas recombinantes de cana-de-açúcar, conhecidas como canacistatinas, CaneCPI-1, CaneCPI-2, CaneCPI-3, CaneCPI-4, CaneCPI-5 e CaneCPI-6, inibidores específicos de cisteína peptidases, produzidas

no laboratório de Biologia Molecular do Professor Dr. Flávio Henrique Silva (Departamento de Genética e Evolução – UFSCar).

No ensaio de inibição também foi utilizado o peptídeo epóxido E-64 inibidor de cisteíno peptidase, conhecido por inibir irreversivelmente as enzimas da família da papaína (BARRET et al., 1982; KIM et al., 1992). Assim, a comparação da atividade proteolítica na ausência e na presença do inibidor forneceu o percentual de inibição que foi calculado pela equação (1). Sendo V_i a velocidade de reação observada na presença de inibidor e V_0 a velocidade de reação observada na ausência do inibidor.

$$\% \text{ Inibição} = 100 \times (1 - V_i/V_0) \quad (1)$$

A inibição da atividade proteolítica foi determinada utilizando o substrato Z-FR-AMC. O procedimento utilizado compreendeu a ativação da enzima por 5 minutos conforme item 3.6.1 e, em seguida, o inibidor foi adicionado e incubado por mais 5 minutos antes de adicionar o substrato. As concentrações dos inibidores foram de 10 nM para cada canacistatina avaliada individualmente e 250 nM para o epóxido E-64 também avaliado individualmente.

Após a triagem, as canacistatinas que apresentaram inibição superior a 90% frente a catepsina L foram selecionadas para determinação da constante de inibição (K_i) do complexo enzima–inibidor. Uma curva de titulação foi feita para cada um dos inibidores de cana-de-açúcar, variando a concentração: 0,625; 1,250; 2,5; 5,0; 10,0 e 20,0 nM. Os resultados foram analisados segundo a equação geral para inibidores competitivos (MORRISON, 1969; COPELAND, 2000), onde o V_i para cada concentração do inibidor foi dividido pela taxa inicial V_0 , e foram ajustados pela equação (2), utilizando o software GraphPad Prism 5.0.

$$\frac{v_i}{v_0} = 1 - \frac{([E] + [I] + K_i^{app}) - \sqrt{([E] + [I] + K_i^{app})^2 - 4[E][I]}}{2[E]} \quad (2)$$

$$K_i^{\text{app}} = K_i \left(1 + \frac{[S]}{K_m} \right) \quad (3)$$

V_i = Velocidade de reação para cada concentração do inibidor;

V_0 = Velocidade de reação na ausência do inibidor;

$[Et]$ = Concentração total da enzima;

$[I]$ = Concentração total do inibidor;

K_i^{app} = Constante de dissociação aparente para o complexo enzima-inibidor;

$[S]$ = Concentração do substrato;

K_m = Constante de Michaelis-Menten.

3.9 Análise da AsCathL por espectrometria de massas

Os experimentos foram realizados pelo Pós Doc. Dr. Hugo César Ramos de Jesus no Laboratório de Bioquímica e Biologia Molecular, UFSCar e no Instituto de Química de São Carlos – USP (Central de Análises Químicas Instrumentais – CAQI).

A espectrometria de massas (MS) é uma técnica capaz de determinar a relação entre massa e carga (m/z) de espécies ionizadas em fase gasosa. Um espectrômetro de massas é um instrumento constituído por uma fonte de íons, analisador de massas, um detector e um sistema de aquisição de dados. Com esses recursos é possível obter informações parciais da sequência primária de proteínas, que podem ser usadas para fins de identificação ou análise de modificações pós-traducionais. Assim, após a etapa de purificação da proteína, o próximo passo é converter a proteína em um conjunto de peptídeos, isso é feito com o uso de enzimas que promovem a clivagem das proteínas em pontos específicos.

Considerando que os MS podem determinar a massa molecular das proteínas intactas, existem inúmeras razões que justificam o uso de peptídeos e não proteínas intactas para a análise. Dentre essas razões, é que as proteínas são difíceis de manusear e degradam-se com facilidade, podendo ainda apresentar problemas de solubilidade. Ademais, a sensibilidade dos MS é consideravelmente menor para os peptídeos e, se o interesse da análise for a identificação das

proteínas, informações inerentes à sequência são necessárias, e nesse sentido, os MS são mais eficientes para obter informação estrutural de peptídeos que possuem até 20 aminoácidos em comparação com a proteína intacta (CANTÚ et al., 2008).

Os peptídeos obtidos são separados por meio de técnicas de cromatografia líquida, ionizados e transferidos para um analisador de massas. Nessa etapa o espectro de massas dos peptídeos oriundo da digestão enzimática é obtido e este resultado indica a relação m/z e, por consequência a massa molecular do peptídeo. Os peptídeos previamente detectados são isolados e submetidos à fragmentação por colisão com moléculas de um gás inerte. O espectro obtido é chamado de espectro de fragmentação ou de MS/MS.

Ao final do processo, os resultados obtidos (MM dos peptídeos e os espectros MS/MS), são usados por softwares de busca para identificar as proteínas nos banco de dados (CANTÚ et al., 2008).

Uma alíquota de 200 μL da enzima purificada (7,6 μM) foi submetida à redução das ligações dissulfeto através da adição de 15 μL (10 nM) de solução de 1,4 DTT (Dithiothreitol) em tampão 50 mM Bicarbonato de amônio. A reação ficou incubada por 30 min em temperatura ambiente. A subsequente alquilação dos grupos tióis livres foi realizada através da adição de 10 μL de 50 mM de solução de IAA (iodoacetamida) em tampão 50 mM Bicarbonato de amônio. A mistura reacional foi mantida a 30 °C por 30 minutos em ambiente com luminosidade reduzida. Por fim, foi realizada a digestão enzimática pela adição de Tripsina porcina (Sequencing Grade Modified Trypsin, Promega) em uma proporção de 1:30 (m/m) enzima: substrato, incubando a reação por 17 horas a 37 °C. Após digestão enzimática, a mistura de peptídeos resultantes foi dessalinizada e fracionada utilizando cartuchos de extração em fase sólida “homemade” C18 StageTips, (RAPPSILBER et al., 2007).

Em seguida, cerca de 50 μg de amostra digerida foram analisadas utilizando-se um sistema cromatográfico em nano fluxo Easy-nLC II (Thermo Fisher Scientific), acoplada a um espectrômetro de massas de alta resolução LTQ Orbitrap Velos (Thermo Fisher Scientific), com uma fonte de ionização em nano elétronspray (Nanospray flex, Thermo Fisher Scientific). A separação cromatográfica dos peptídeos foi realizada utilizando-se uma coluna de fase reversa C18-A2 Thermo Easy (10 cm L x 75 μm D.I., 3 μm , 120 Å, Thermo Fisher Scientific) e uma pré-coluna C18-A1 Thermo Easy (2 cm L x 75 μm D.I., 3 μm , 120 Å, Thermo Fisher Scientific).

As amostras foram eluídas durante 150 minutos, com um gradiente de 5% de solvente B (Acetonitrila, contendo 0,1 % de ácido fórmico) e 95 % de solvente A (água deionizada, com 0,1 % de ácido fórmico) durante 5 minutos, de 5-40 % de B, no intervalo de tempo de 5 a 100 minutos; de 40-99% de B (de 100 a 115 minutos); 99% de B, (de 115 a 135 minutos); de 99-5 % de B (de 135 a 145 minutos) e de 5% de B (de 145 a 150 minutos), a um fluxo de 200 nL/min.

As análises foram feitas com fragmentação por dissociação induzida por colisão (CID, *Collision-induced dissociation*), com o espectrômetro de massas no modo full-scan, no intervalo de 300 a 1800 m/z, em modo positivo, a voltagem do nanospray de 2.3 KV e temperatura da fonte em 250 °C. Os dados foram adquiridos utilizando o *software* XCalibur, versão 2.0.7 (Thermo Fisher Scientific).

As corridas de LC-MS foram processadas (deconvolução, deisotopização e correção de medida de m/z) pelo *software* MASCOT Distiller (Matrix Science Ltd) e analisadas por busca contra banco de dados empregando-se o sistema MASCOT v.2.4. (Matrix Science Ltd). Os parâmetros da busca foram digestão com tripsina com até 2 sítios de clivagem perdidos, carbamidometilcisteínas como modificação fixa, oxidação em resíduos de metionina como modificação variável e tolerância de massa de peptídeos e fragmentos ambos em $\pm 0,15$ Da. As buscas foram feitas no banco de dados NCBI nr.

3.10 Análise das sequências, alinhamentos e filogenia

A construção da árvore filogenética usando o programa R foi feita com a ajuda do aluno de doutorado em Bioinformática Amr Galal do IME - Instituto de Matemática e Estatística - USP – Ribeirão Preto.

Para o alinhamento, a predita sequência de aminoácidos da catepsina L da formiga cortadeira *A. sexdens* foi comparada com sequências de formigas depositadas no banco de dados global NCBI (<http://www.ncbi.nlm.nih.gov>), utilizando o Blastp. Em paralelo, uma busca por sequências de aminoácidos das catepsinas L de variadas ordens (Coleoptera, Diptera, Hemíptera e Lepdoptera) de insetos também foi realizado. Assim, após as identificações do conjunto de proteínas homólogas feitas para as duas análises, as sequências foram alinhadas em todo o seu comprimento, a fim de obter um alinhamento de sequência múltipla (MSA), utilizando o *software* Clustaw (LARKIN et al., 2007).

Para a construção da árvore filogenética foi usado o programa estatístico R disponibilizado na plataforma R, ferramenta utilizada em bioinformática (R CORE TEAM, 2017). O pacote R *msa* fornece uma interface unificada para os métodos de alinhamento de sequências múltiplas mais comuns (BODENHOFER et al., 2015). A construção de árvores filogenéticas utilizadas na plataforma R depende do alinhamento feito no ClustalW e da utilização do pacote *ape* para converter os dados em árvore filogenética numérica, e do pacote *ggree* para visualizar o alinhamento da árvore e das cores dos aminoácidos (YU et al., 2017). As árvores filogenéticas foram construídas pelo método de *Neighbor Joining* criado por SAITOU e NEI (1987), e o valor de *bootstrap*, $n = 100$.

4 RESULTADOS E DISCUSSÃO

4.1 Amplificação do DNA referente à catepsina de *A. sexdens*

De posse da sequência de nucleotídeos depositada de catepsina L da formiga *Acromyrmex echinator*, o oligonucleotídeo *sense* foi desenhado retirando os primeiros 54 resíduos de nucleotídeos da sequência de DNA depositada, o que levou à eliminação dos primeiros 18 resíduos de aminoácidos da região N-terminal correspondente ao peptídeo sinal. A Figura 4.1 exibe a tradução da sequência completa de DNA (complete CDS), em sequência de aminoácidos da catepsina L de *A. echinator*. A sequência em vermelho sugere que seja o peptídeo sinal, a sequência em azul o pró-peptídeo e em preto a enzima madura. O oligonucleotídeo *reverse* também foi desenhado baseado na sequência de nucleotídeos *A. echinator*.

FIGURA 4.1- Tradução da sequência completa de DNA (completo CDS) em sequência de aminoácidos da catepsina L de *A. echinator*

MKSVVLLLLAVFTTGQAISFNKILDAEWFIFKTHHKKIYKSSVEEGYRMKIFLDNKRKIA
 EHNRYELNEVPYKLG MNKYGDMLHHEFVNTLNGFNKSEKAQKQFMGATFISPANVELP
 KEVDWRKHGAVTEVKDQGHCGSCWAFSTTGSLEGQHFRQTGILVSLSEQNLIDCSGNYG
 NEGCGGLMDNAFKYVRDNKGLDTEKSYPYEAENDKCRYNPRNSGAIDTGFVDIPRGNE
 HKLKAAVATIGPVSVAIDASHESFQLYSEG VYFDPECDSDNLDHGVLIVGYGTD SKTGHDY
 WLGKKS YIKMARNKDNHCGIASSASYPLV

Fonte : Própria

Nos desenhos dos oligonucleotídeos foram incluídos sítios de clivagem para enzimas de restrição *NcoI* (Invitrogen™) para o oligonucleotídeo *sense*, e *EcoRI* (Invitrogen™) para o oligonucleotídeo *reverse*, os quais permitiriam a clonagem do gene no vetor de expressão pET-32a(+) (Novagen).

A escolha dos sítios de restrição foi baseada na ausência desses sítios no gene de interesse, mas que são existentes no vetor de expressão escolhido. A Figura 4.2 apresenta as sequências dos oligonucleotídeos sintetizados para amplificação do gene de interesse sem o peptídeo sinal. Os sítios de restrição sublinhados correspondem às enzimas *NcoI* e *EcoRI*, respectivamente. As setas vermelhas indicam o ponto de clivagem das enzimas.

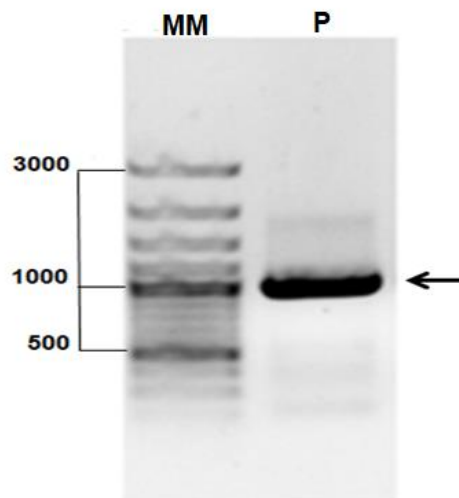
FIGURA 4.2- Oligonucleotídeos sintetizados para a amplificação da ORF CathL

	Vetor de Expressão	Sequência de Oligonucleotídeos
<i>Sense</i> (5'-3') (<i>Nco</i> I)	pET-32 a (+).	5'GCCATGGCCATTTTCCTTCAACAAGATTTTA3'
<i>Reverse</i> (5'-3') (<i>Eco</i> RI)	pET-32 a (+).	5'CGCGAATTCTTAAACGAGAGGATAGCTGGC3'

Fonte: Própria

O cDNA, sintetizado a partir RNA total extraído da formiga *A. sexdens* (formiga adulto), foi utilizado como molde na reação de PCR. Desse modo, a obtenção do inserto utilizando a reação de amplificação do DNA promovido por PCR foi realizada obtendo a sequência CathL, ORF que possui a região codificadora para a síntese da catepsina L da *A. sexdens*. Na Figura 4.3 têm-se o resultado da amplificação.

FIGURA 4.3- Análise em gel de agarose 1% da amplificação da ORF



Fonte: Própria

A Figura acima exhibe: **MM**: Marcador de massa molecular Gene Ruler™ 100 bp Plus DNA Ladder (Therm Scientific); **P**: Produto da amplificação utilizando a Pfu DNA polimerase e cDNA como molde na reação de PCR. O fragmento de DNA amplificado com massa molecular aparente em torno de 1000 bp bem próximo da sequência predita depositada (NCBI) da catepsinas L da formiga cortadeira *A. echinator* com 935 pb, excluindo a região do peptídeo sinal.

A sequência de DNA obtida após análise do sequenciamento apresentou 91% de identidade com o gene da formiga cortadeira *A.echinator*. Um alinhamento da sequência dos nucleotídeos da formiga *A.echinator* com a sequência do produto de PCR obtido (CathL) pode ser visualizado Figura 4.4.

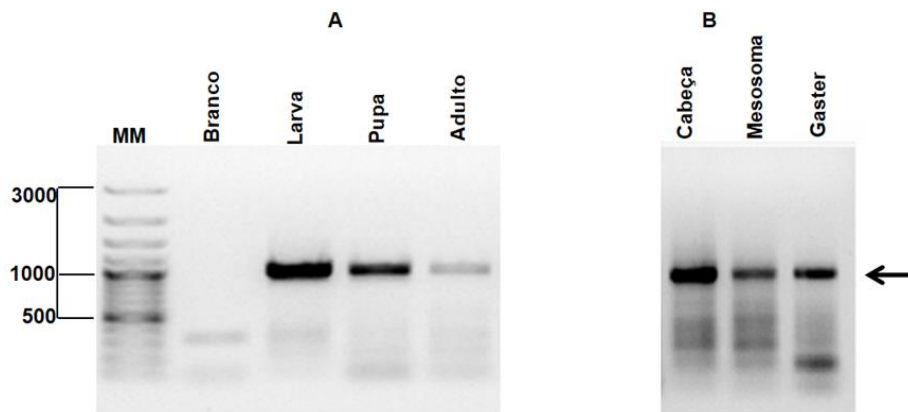
FIGURA 4.3- Alinhamento das sequências de nucleotídeos de uma catesina da *A.echinator* com a sequência do gene obtido CathL. A sequência sombreada em cinza é a região do peptídeo sinal da *A.echinator*, ausente na sequência predita do CathL. (*) resíduos idênticos em ambas as sequências



Fonte: Alinhamento feito no ClustalW

A expressão da ORF CathL foi analisada nos estágios de desenvolvimento do inseto (larva, pupa, adulto) e em partes do corpo da formiga: cabeça, tórax e gaster. Os resultados podem ser visualizados na Figura 4.5.

FIGURA 4.4 - Análise qualitativa da expressão da ORF CathL em gel de agarose 1%. Em **A** - Estágios do desenvolvimento - **B** - Partes do corpo de *A.sexdens*. **MM**: Marcador de massa molecular Gene Ruler™ 100 bp Plus DNA Ladder (Thermo Scientific).



Fonte: Própria

O perfil de expressão sugere que a ORF CathL codificante para catepsina L de *A.sexdens* está presente em cada fase da vida do inseto. Existem diversos estudos descritos na literatura sobre atuação de catepsina L durante a metamorfose, na reconstrução dos tecidos do inseto (HEGEDUS et al., 2002). Em *Helicoverpa armigera*, a catepsina L por exemplo, está envolvida na degradação da cutícula durante a muda no inseto (LIU et al., 2006). Segundo HOMMA et al., (1994), a ação proteolítica da catepsina L na remodelação dos tecidos em insetos está relacionada à degradação da membrana basal por proteólise limitada. A catepsina L em insetos é considerada uma enzima digestiva (TERRA et al., 1996), porém, como nos outros animais, espera-se que a catepsina L ocorra no ambiente lisossomal (BARRET et al., 1998).

Nessa análise preliminar, não há como inferir qual seria a função da enzima (digestiva, lisossomal, ou estar envolvida em outras funções). O uso da técnica da imunolocalização utilizando anticorpos específicos para localização celular e experimentos de PCR quantitativo em tempo real para detecção de onde a ORF é mais expressa (análise quantitativa), podem contribuir para compreensão da função biológica da enzima na fisiologia do inseto.

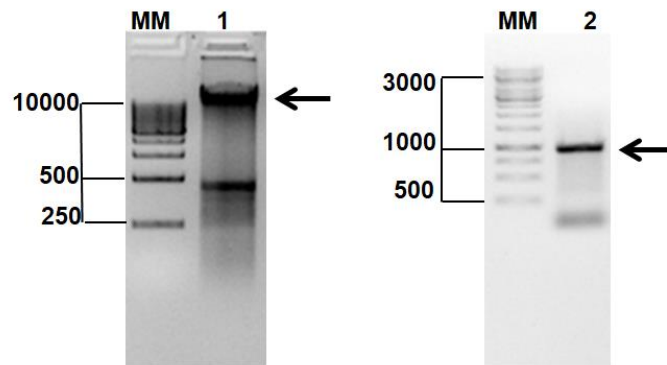
Outra técnica para detectar e quantificar proteólise pode ser feita com uso de sondas moleculares ligadas à enzima alvo. A avaliação da liberação do sinal fluorescente após reações químicas em sondas fluorogênicas desencadeadas pela ação enzimática é uma estratégia bastante utilizada para o monitoramento da atividade proteolítica. Essas sondas ativadas fornecem uma leitura indireta da atividade da enzima que pode ser monitorada *in vitro* e *in vivo* usando métodos bioquímicos, espectrometria de massa e microscopia de fluorescência. Usando marcadores é possível determinar quando e onde as enzimas são ativadas durante um determinado estímulo celular (BLUM et al., 2007; SANMAN e BOGYO, 2014).

A relação simbiótica entre a formiga *A.sexdens* e o fungo *L.gongylophorus*, lançou dúvida sobre a origem da enzima. Visto que a formiga se alimenta desse micro-organismo, não era possível afirmar se a enzima era do organismo da formiga ou se era do fungo. Assim, foi feita a extração do DNA genômico da cabeça da formiga, parte do corpo que poderia estar livre de contaminação pelo fungo. O DNA genômico foi extraído, e o valor quantificado foi de 430 ng/μL. Na Figura 4.6 têm-se o resultado da amplificação da ORF tendo como molde o DNA genômico. A seta preta indica a ORF amplificada partindo de DNA genômico como molécula molde na PCR, com aproximadamente 1000 pb, assim como o DNA amplificado a partir de cDNA, o que sugere não haver íntrons na sequência do gene de catepsina. O resultado obtido afirma que a ORF de estudo é da formiga *A.sexdens* e não do fungo *L.gongylophorus*.

O genoma da *A.sexdens* não está sequenciado, mas da formiga *A.cephalotes* está, e o gene que codifica para catepsina L dessa formiga depositado no NCBI (*Reference Sequence*: NW_012130065.1) possui 7197 pb (íntrons e exons). Em um alinhamento feito no Blastn foi possível visualizar os nucleotídeos do gene da catepsina extraído de *A.sexdens* pareando com os exons do gene da *A.cephalotes*, com uma identidade de 80%.

FIGURA 4.5 - Análise em gel de agarose 1% do DNA genômico extraído da cabeça da formiga *A.sexdens* e amplificação da ORF a partir do DNA genômico.

MM: marcador de massa molecular 1Kb DNA ladder (Thermo Scientific); **1:** DNA genômico; **2:** Amplificação da região codificante da catepsina L na formiga *A.sexdens* a partir do DNA genômico.

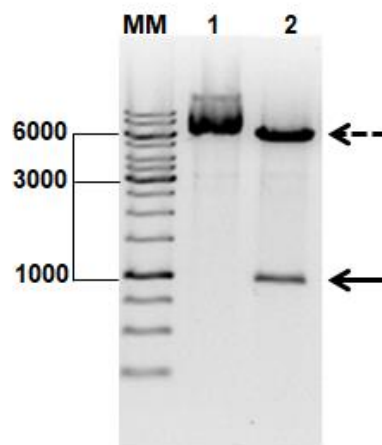


Fonte: Própria

4.2 Clonagem da ORF (CathL) em vetor de expressão e análise da sequência

O clone positivo foi analisado por restrição, utilizando as enzimas *NcoI* e *EcoRI* sob as condições descritas na Tabela 3.4. A Figura 4.8 exibe o resultado da análise da reação de restrição do clone.

FIGURA 4.6- Análise em gel de agarose 1% da análise de restrição utilizando as enzimas de restrição *NcoI* e *EcoRI*. **MM:** marcador de peso molecular 1Kb DNA ladder (Thermo Scientific); **1:** O clone denominado AsCathL; **2:** O produto da clivagem gerou dois fragmentos de DNA, um que apresenta uma banda de aproximadamente 6000 pb referente ao vetor de expressão (seta preta tracejada) e o outro fragmento de 1000 pb referente à banda da ORF CathL (seta preta).



Fonte: Própria

A escolha do vetor de expressão pET32-a(+) foi estratégico, uma vez que ele permite a expressão da proteína de interesse em fusão (na região N terminal) com seis resíduos de histidina (His.tag), sequência de reconhecimento para detecção e purificação da enzima. A região dos resíduos de histidina é precedida pela região Trx-Tag. A tirredoxina (Trx) é uma enzima que sofre reação redox através da oxidação reversível de dois grupos tiol cisteínico a um dissulfeto resultando na transferência de dois elétrons e dois prótons. A fusão de proteínas com Trx.tag leva a uma maior formação de ligações dissulfeto, diminuindo a formação de agregados de proteínas WONG S. H. (2005). Nessa construção, optou-se pelo uso do *stop códon* da ORF, de modo que a proteína não teria uma cauda de histidina C-terminal, e sim no N-terminal.

O resultado do sequenciamento do clone AsCathL mostrou que a ORF correspondente à catepsina L está em fase correta de leitura. A predição da estrutura primária da enzima foi realizada a partir da tradução da sequência de nucleotídeo utilizando o software online *translate tool* <https://web.expasy.org/translate/>. A predita pró-enzima é composta por 319 aminoácidos, com massa molecular teórica estimada em torno de 35 kDa

O alinhamento entre as sequências de catepsinas L depositadas no NCBI e a predita catepsina L de *A.sexdens* (trabalho), foi realizado com o objetivo de buscar identidade entre as sequências. As sequências utilizadas foram: da formiga (*Atta cephalotes*) (XP_012059550), da vespa (*Delia coarctata*) (ACR56863) e a sequência predita para catepsina L da formiga (*Acromyrmex echinator*)(EGL69825). A sequência da *A.Sexdens* apresentou 91% de similaridade com a *A.echinator*, 94,04% com *A.cephalotes* e 67,92% com a *D.coarctata*. Observou-se uma sequência consenso entre as catepsinas L e a de trabalho nos aminoácidos catalíticos: cisteína (C), histidina (H) e asparagina (N), destacados em vermelho nas sequências avaliadas. A região sombreada em cinza corresponde à sequência do peptídeo sinal, não presente na sequência da AsCathL da *A. sexdens*. A sequência do predito pró-domínio está representada em azul, contendo os aminoácidos responsáveis pela inibição, representados em cor azul, sombreados na cor cinza, a sequência predita da proteína “madura” está em preto. A análise do alinhamento pode ser visualizada na Figura 4.9.

FIGURA 4.7- Alinhamento das sequências preditas para catepsinas L da vespa *D. coarctata*, da catepsina L *A. Sexdens* (trabalho), da formiga *A. cephalotes* e de *A. echinator*

```

CatLcoarctata  MRTV--LALLALVAFVQAISITDVIKEEWQTFKMEHRKNYLSEVEERFRMKIENENRHKI
CatLechinatior  MKSVVLLLLAVFTTQQAISFNKILDAEWFIFKTHHKKIYKSSVEEGYRMKIFLDNKRKI
CatLSexdens      -----ISFNKILDVEWFSFKTNHKKIYKSSVEEGLRMKIFLNNKHKI
CatLcephalotes  MKSVVLLLLAVFTTQQAISFNKILDIEWFSFMTNHHKKIYKSSIEEGYRMKIFLDNKRKI
                **:. . . . . ** * .:* * * .:** ***** :*:.**

CatLcoarctata  AKHNQLYAQGKVSFKLGLNKYADMLHHEFKETMNGYNHTMRKELRAQEGFNGITYIS PAN
CatLechinatior  AEHNRYKYLNEVPYKLGMNKYGDMHLHEFVNTLNGFNKS----EKAQKQFMGATFIS PAN
CatLSexdens      IEHNRYKYLNEVPYKLGMNKYSDMLHHEFVNTLNGFNKS----VKAQKQFMGATFIS PAN
CatLcephalotes  VEHNRYKYLNEVPYKLGMNKYCDMFHHEFVNTLNGFNKS----VKAHKQFMGATFIS PAN
                :*: * .:* :***:*** **:***** :*:**:*: :*: * * *:*****

CatLcoarctata  VQVPKAVDWRQHGAVTSVKDGHCSCWFSSTGSLEGQHFRKAGVLVSLSEQNLVDCST
CatLechinatior  VELPKEVDWRKHGAVTEVKDQGHCSWAFSTTGSLEGQHFRQTGILVSLSEQNLIDCSG
CatLSexdens      VELPKEVDWRKHGAVTEVKDQGHCSWAFSTTGSLEGQHFRQTGILVSLSEQNLIDCSG
CatLcephalotes  VELPKEVDWRKHGAVTEVKDQGHCSWAFSTTGSLEGQHFRQTGILVSLSEQNLVDCSG
                *:.** *****:*****.*****:***:*****:*.*****:*.*****:***

CatLcoarctata  KYGNNGCNGGLMDNAFRYIKDNGGVDTEKSYPYEGIDDSCHFENKATVGATDTGFVDIPQG
CatLechinatior  NYGNEGCNGGLMDNAFKYVRDNKGLDTEKSYPYEAENDKCRYNPRNSGAIDTGFVDIPRG
CatLSexdens      KYGNEGCNGGLMDNAFKYVRDNKGLDTEKSYPYEAEDKCRYNPRNSGATDVGFVDIPRG
CatLcephalotes  NYGNQGCNGGLMDNAFKYVRDNKGLDTEKSYPYEAENDKCRYNPRNSGATDVGFVDIPRG
                :***:*****:***:*.** *:*****. .*.**:* . ** *.*****:*

CatLcoarctata  DEEAMKAVATMGPVAVAIASNESFQLYSEGVYNDPNCSSDNLDHGVLVVGYGTD-KDG
CatLechinatior  NEHKLKAAVATIGPVSVAIDASHESFQLYSEGVYFDPECDSDNLDHGVLIVGYGTDKSTG
CatLSexdens      NEEKLKAAVATIGPVSVAIDASHESFQFYSEGVYFDPECNTENLDHGVLIVGYGTDKSTG
CatLcephalotes  NEQKLKAAVATIGPVSVAIDASHESFQFYSEGVYFDPECDSENLDHGVLIVGYGTDKSTG
                :* . : *****:***:*****:***:***** **:* .:*****:***** * *

CatLcoarctata  QDYWLVKNSWGTWGDQGYIKMARNQDNQCGIATASSFPTV-
CatLechinatior  HDYWLG-----KKSIIKMARKDNHCGIASSASYPLV-
CatLSexdens      HDYWLVKNSWGETWGEKGYIKMARNKNHCGIASSASYPLV-
CatLcephalotes  HDYWLVKNSWGETWGEKGYIKIARNKNHCGIASSASYPLVS
                :**** . . . ***:***:*.*****:*.** *

```

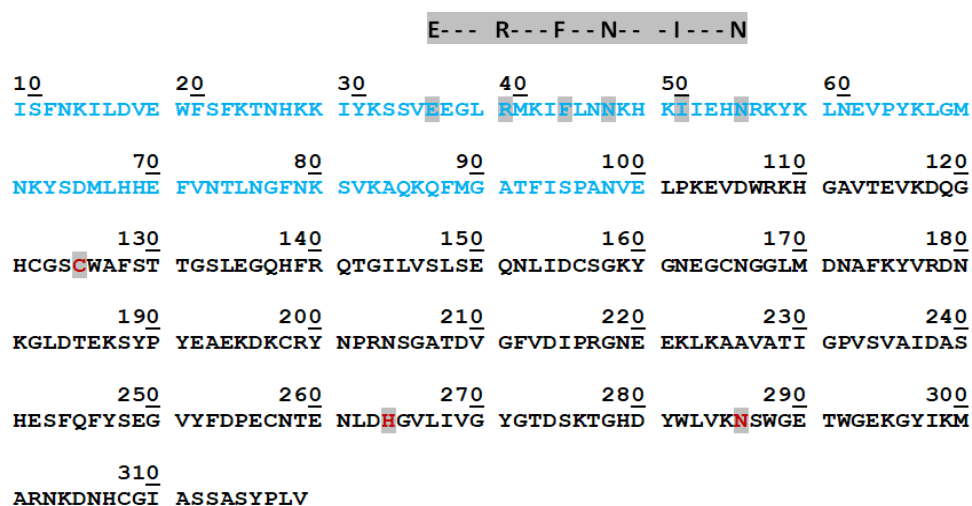
Fonte: Alinhamento feito no programa ClustalW. (*) resíduos idênticos em ambas as sequências, (:) resíduos de alta similaridade nas sequências e (.) resíduos com similaridade média.

A tríade catalítica das preditas sequências acima são características da cisteína peptidases (EC 3.4.22), família das enzimas proteolíticas que contêm uma cisteína no sítio ativo. A catálise prossegue através de um intermediário tioéster e é facilitada por uma cadeia lateral de histidina próxima, e uma asparagina completando a tríade catalítica. As sequências em volta da tríade catalítica são bem conservadas como observado nas sequências acima analisadas. Esta tríade catalítica é detectada principalmente na família C1 das proteases *papain-like*.

A AsCathL foi expressa na forma de zimogênio, um precursor inativo, com uma região pró-peptídica contendo o motivo conservado $E^R F^N I^N$ ^{27,31,35,38,42,46}

(Glu27, Arg31, Phe35, Asn38, Ile42 e Asn 46), aminoácidos envolvidos na inibição da enzima, característico das enzimas *papain-like*. Este pró-peptídeo deve ser removido para formar a enzima madura ativa (SANMAN e BOGYO, 2014). A sequência do pró-peptídeo no domínio da alfa hélice interage com a enzima ao longo do sítio ativo da enzima madura numa conformação predominantemente estendida na direção oposta da ligação com o substrato, bloqueando o sítio ativo (TURK e GUNCAR 2003). O pro-peptídeo, de fato, é um inibidor da enzima madura como demonstrado por dados cinéticos (TURK e GUNCAR 2003; GUAY et al., 2000), porém possui um papel importante na proteína agindo como uma chaperona intramolecular promovendo o correto enovelamento da proteína (YAMAMOTO *et al.*, 1999). Na Figura 4.9 ilustra a sequência apresentada ausente do peptídeo sinal na região N-terminal. A sequência marcada em azul evidencia a predita região pró-peptídica que se estende até a posição 100. O motivo inibitório conservado da pró-peptídeo **E²⁷R³¹F³⁵N^{38,42}I⁴⁶** está destacado na cor azul com caixa na cor cinza na região pró-peptídica. A predita enzima madura se estende do resíduo 101 ao 319. Os resíduos conservados da tríade catalítica estão destacados em vermelho com sombreamento em cinza.

FIGURA 4.8 - Sequência de aminoácidos da AsCathL obtida pela tradução da sequência de nucleotídeos

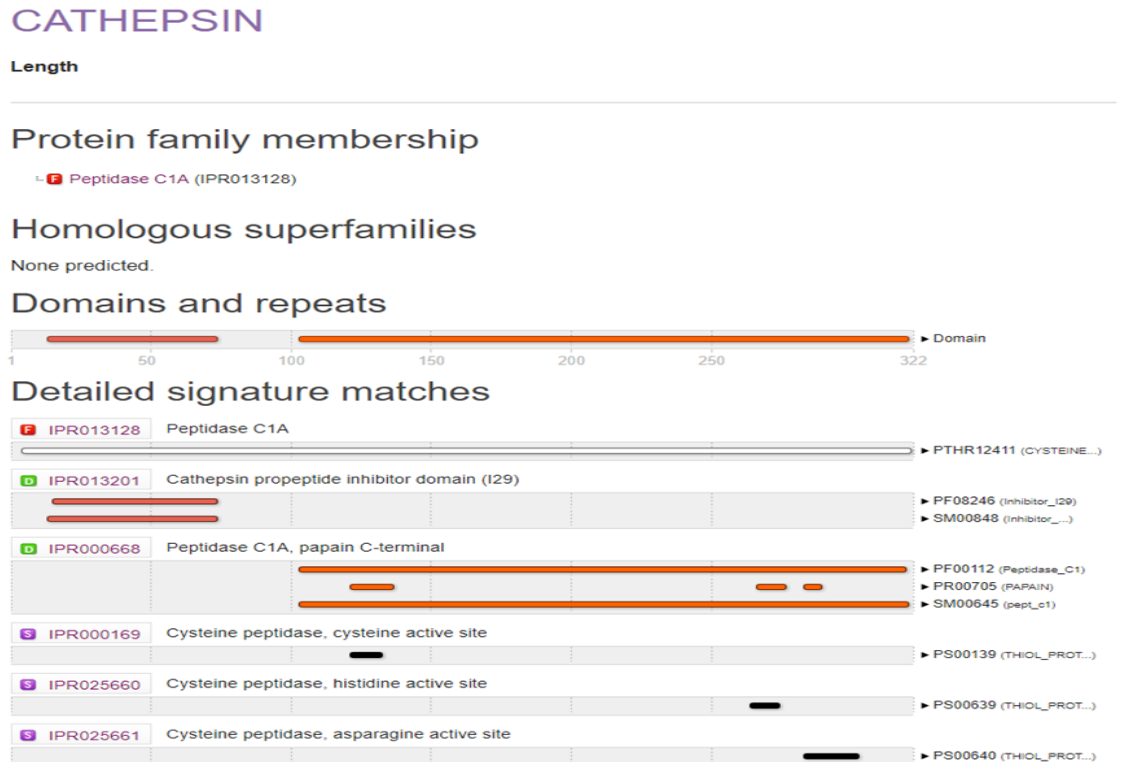


Fonte: Própria

A sequência foi submetida à análise de classificação, utilizando o *database on line* Interpro. A sequência foi classificada como uma catepsina, uma peptidase pertencente à família C1 *papain-like*, com a região inibitória e a presença

da tríade catalítica (Cisteína, Histidina e Asparagina). A Figura 4.11 detalha o local da região pró-peptídica, a região catalítica e a região da enzima madura.

FIGURA 4.9- Análise da sequência e classificação



Fonte: *database on line Interpro*. - <https://www.ebi.ac.uk/interpro/>

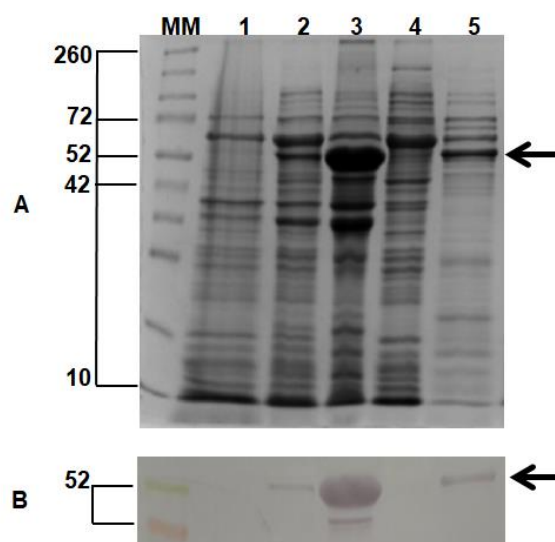
4.3 Expressão da proteína recombinante em *E. coli*

Para obtenção da pró-enzima, o DNA plasmidial contendo a ORF de interesse foi transformado em linhagem de *E. coli ArcticExpress* (DE3)RIL, assim a colônia transformante selecionada foi inoculada em LB ágar com os antibióticos. Para a expressão da AsCathL, experimentos variando a concentração do IPTG (0,1; 0,2; 0,3; 0,4; 0,5 e 1.0 mM) foram necessários, com o objetivo de ter mais proteína na fração solúvel. Porém, na análise qualitativa em gel SDS, notou-se que quanto maior a concentração do indutor, maior era quantidade de proteína na fração insolúvel. Uma vez que o sistema de *E. coli* proporciona rápido crescimento, acoplado a uma concentração alta de indutor, há a hipótese de que com a produção acentuada da proteína, esta não teria o tempo hábil de enovelamento, precipitando (BANEYX, 1999).

Após os ensaios experimentais variando a concentração do indutor, optou-se pela concentração de 0,2 mM na indução da expressão da proteína. Para a lise celular, as células foram ressuspensas em um tampão alcalino (50 mM Tris, pH 8,0), e posterior purificação em coluna de níquel foi realizada para obtenção da pró-enzima. A análise em SDS-PAGE revelou uma proteína parcialmente purificada com massa molecular aparente de 52 kDa, confirmada em *western blotting*. A massa com peso molecular maior do que a esperada, compreende as regiões Trx.Tag e His.Tag componentes do plasmídeo pET32a+, somados aos 36 kDa peso molecular teórico da pró-enzima. A pró-enzima foi expressa e purificada inativa em pH alcalino (Tris-HCl, pH 8.0) para evitar a atividade auto-catalítica de ativação do zimogênio e subsequente auto-degradação da enzima madura, características de cisteíno peptidases em ambiente ácido. A auto-ativação de catepsinas lisossomais, por exemplo, pode ser evitada na presença de pH maior que 6.0 (BROMMER et al., 2004), pois o pH alcalino parece ser essencial para a formação de dissulfetos (TOBBELL et al., 2002). O gel da expressão, purificação em coluna de níquel e o resultado do *western blot* podem ser visualizados na Figura 4.12.

FIGURA 4.10 - **A**- Análise em SDS-PAGE. **B**- *Western blot* (anti-His).

A- A expressão e purificação da pro-enzima em coluna de Ni-NTA. **B**-*Western blot* (anti-His). **MM**: Marcador de massa molecular SpectraMulticolor Broad Range Protein Ladder - Thermo Scientific. **1**: Cultura não induzida; **2**: Cultura induzida por 24 horas com 0.2 mM IPTG; **3**: *Pellet* após sonicação; **4**: Sobrenadante; **5**: Fração eluída com 250 mM Imidazol.



Fonte : Própria

As células de *E.coli arcticExpress* fornecem uma abordagem para aumentar o rendimento da proteína heteróloga na fração solúvel utilizando baixas

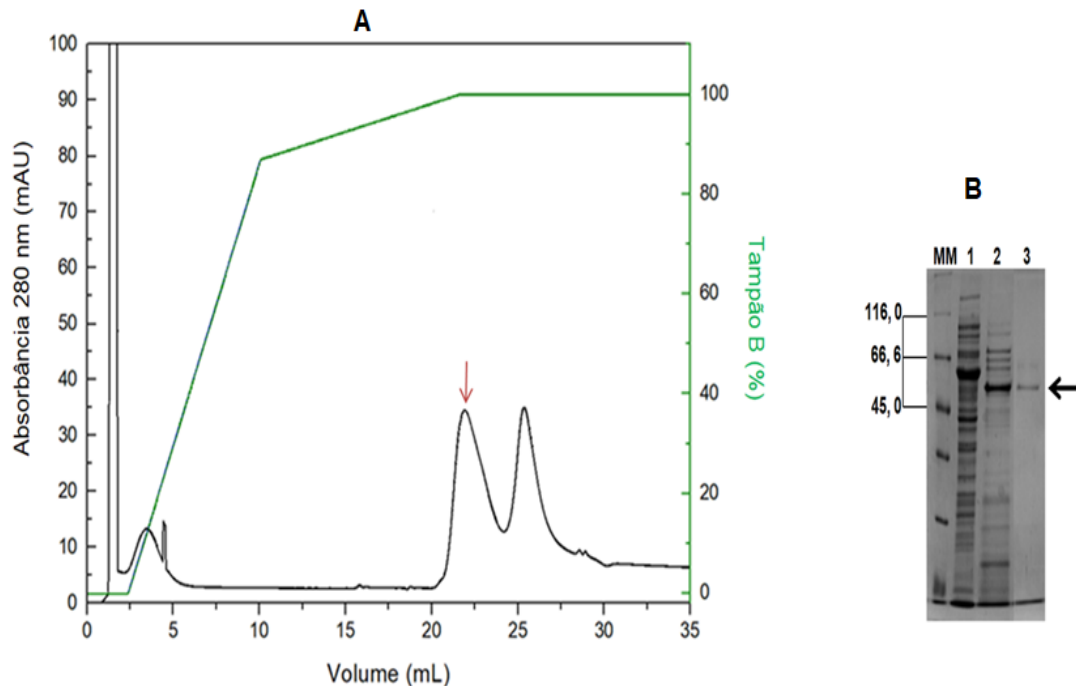
temperaturas de expressão. Essas células co-expressam as chaperoninas Cpn10 e Cpn60 originárias da bactéria *Oleispira Antarctica* e têm 74% e 54% de identidade de aminoácidos com as chaperonas GroEL e GroES de *E.coli* respectivamente, e mostram alta desempenho no enovelamento das proteínas a temperaturas de 4-12°C (FERRER et al., 2003). Quando expressas em células *ArcticExpress*, essas chaperonas proporcionam um melhor processamento de proteínas a temperaturas mais baixas aumentando potencialmente o rendimento de proteína recombinante solúvel e ativa. Na expressão de uma proteína recombinante, essas chaperonas são expressas constitutivamente e são facilmente detectadas em SDS PAGE. A Cpn60 pode ser observada em gel com massa molecular em torno de 56,900 Da.

Embora haja exemplos de expressão bem sucedida de proteínas recombinantes usando as cepas de *arcticExpress* (JOSEPH et al., 2008; ANNAMALAI et al., 2009), neste trabalho, foi possível obter uma pequena quantidade de proteína na fração solúvel, porém com ação enzimática suficiente para ser detectada utilizando o substrato fluorogênico Z-FR-MCA.

A primeira etapa de purificação utilizando resina de níquel proporcionou uma parcial purificação da proteína, assim uma segunda etapa de purificação foi adicionada ao processo a fim de se obter uma proteína com um grau de pureza mais elevado. A amostra eluída da coluna de níquel foi submetida à segunda etapa de purificação fazendo uso da cromatografia de troca iônica, *Hitrap QFF* (resina aniônica), acoplada ao sistema FPLC utilizando o equipamento AKTA. As frações eluídas da coluna foram coletadas em frações de 1 mL e foram submetidas à atividade enzimática. A fração que apresentou atividade foi então concentrada 4 vezes utilizando o sistema de ultracentrifugação (Amicon, membrana de corte 10 kDa). O perfil cromatográfico e a análise em SDS-PAGE da fração eluída da coluna com atividade podem ser visualizados na Figura 4.13.

FIGURA 4.11 - Perfil cromatográfico (A) e análise em SDS-PAGE da cromatografia em coluna QFF (B).

A - Perfil cromatográfico da purificação da pro-enzima em resina aniônica QFF (Q Sepharose Fast Flow) equilibrada com 50 mM Tris- HCl, pH 8.0 (Tampão A). A proteína foi eluída com 87% de 50 mM Tris-HCl, 1 M de NaCl, pH 8.0, (Tampão B). **B** - Análise em SDS-PAGE da fração eluída da coluna com atividade. MM: Marcador de massa molecular (Thermo Unstained Protein MW Marker - Thermo Scientific); 1: Fração solúvel (sobrenadante); 2: Fração eluída do NI-NTA; 3: Fração 22 (QFF) eluída da coluna e com atividade enzimática.



Fonte: Programa OriginPro 8.0 ; Gel SDS-PAGE: Elaboração própria

Os resultados obtidos da expressão e purificação da pró-enzima mostraram que o sistema de *E.coli*, mesmo com engenharia modificada para crescimento de cultura a baixas temperaturas, não foi eficiente para expressar catepsinas na forma solúvel em quantidade alta, pois a maioria das proteínas produzidas estava em corpos de inclusão.

A superexpressão das proteínas pode ser um dos fatores de formação de corpos de inclusão. Proteínas recombinantes expressas em *E. coli* enfrentam frequentemente o problema de solubilidade (BANEYX, 1999), uma vez que as ligações de dissulfeto, essenciais para o correto enovelamento da proteína, não conseguem se formar a tempo, ou se ligam de maneira indesejável, favorecendo a formação de agregados insolúveis de proteínas, levando a perda da atividade biológica da enzima (WONG, 2005).

4.4 Recuperação da AsCathL a partir dos corpos de inclusão

Os corpos de inclusão contêm proteínas recombinantes relativamente puras, o que estimula o desenvolvimento de técnicas que auxiliem na recuperação do correto enovelamento da proteína, na sua forma biologicamente ativa. Embora diversos protocolos estejam descritos na literatura, a estratégia geral é baseada em três passos: isolamento e lavagem dos corpos de inclusão, solubilização das proteínas agregadas e *refolding*. Num procedimento típico, as formas agregadas são desnaturadas em alta concentração de ureia e/ou cloreto de guanidínio (GdnHCl). Esses reagentes químicos são usados para diminuir as interações não covalentes entre moléculas de proteínas. (YAMAGUCHI et al., 2014).

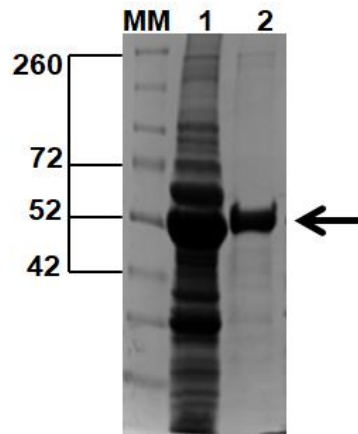
O re-enovelamento de proteínas desnaturadas para proteínas ativas ocorre pela remoção do desnaturante, que pode ser feito pelo uso da diluição da proteína recombinante (CLARK, 2001), diálise ou cromatografia de gel filtração (VILLAVARDE et al., 2003). Alguns aminoácidos como arginina e seus derivados são alguns dos aditivos químicos adicionados ao tampão de *refolding* (SHENAI et al., 2000). Esses aminoácidos são utilizados como inibidores da agregação de proteínas. Este tipo de recurso, juntamente com a adição do glicerol (estabilizador de proteína) e da ação do par redox glutationa redutase/oxidase, têm sido utilizados no processo do re-enovelamento, aumentando os rendimentos do processo e diminuindo a agregação durante essa etapa. A eficiência do processo (rendimento) da proteína é estimada pela atividade enzimática.

A etapa de purificação anterior ao *refolding* foi uma estratégia encontrada para a obtenção de maiores rendimentos. A purificação por cromatografia de afinidade (NI-NTA) proporcionou tal rendimento uma vez que a construção foi feita com um His-tag. Neste trabalho, apenas um etapa de diluição da proteína após a sua solubilização foi necessária para obter pró-enzima re-novelada. Para evitar a auto-ativação, a pró-enzima foi estocada em tampão alcalino, Tris HCl pH 8,0 à temperatura 20 °C. Utilizando essa técnica, foram obtidos cerca de 5 mg de proteína / L de meio de cultura. O protocolo seguido (SHENAI et al., 2000) mostrou ser de fácil execução, pois apenas algumas alterações no protocolo foram necessárias para que fosse possível obter a proteína biologicamente ativa, e todo o processo mostrou ser reprodutível. A atividade enzimática foi o

parâmetro mais conclusivo para o sucesso do processo. Na Figura 4.14, é exibida a proteína recuperada a partir dos corpos de inclusão analisada por SDS-PAGE.

FIGURA 4.12- Análise em SDS-PAGE da proteína recuperada a partir dos corpos de inclusão, utilizando o *refolding*.

MM: Marcador de massa molecular; **1:** *pellet* contendo a proteína precipitada; **2:** pró-enzima obtida com peso molecular de 52 kDa.



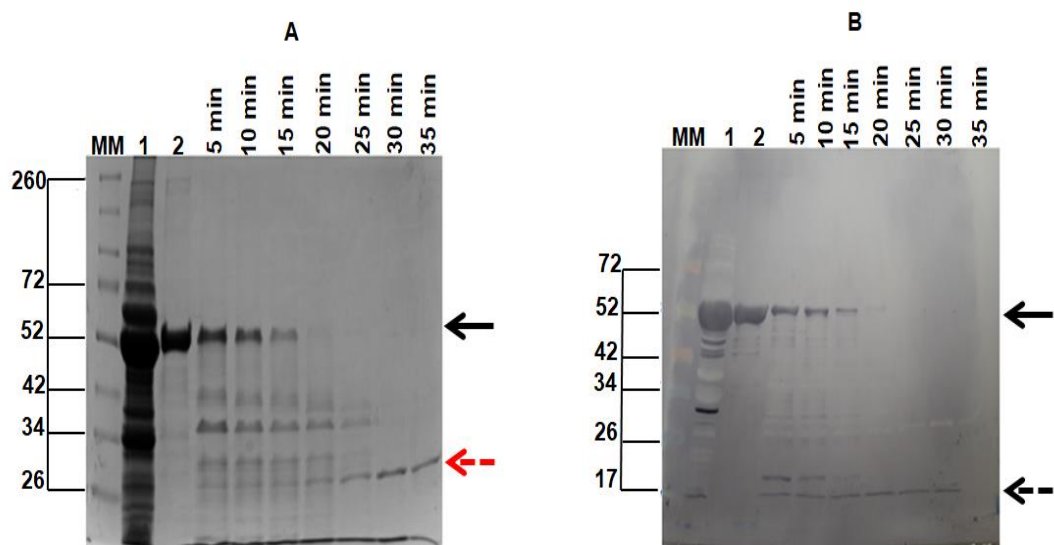
Fonte: Própria

Para gerar uma peptidase proteoliticamente ativa, a região pró-peptídica precisa ser removida, uma vez que a cisteína peptidase foi expressa na forma de zimogênio. A remoção acontece por autocatálise incubando o precursor da proteína madura em tampão ácido. A auto-ativação da pró-enzima para enzima ativa acontece logo após os primeiros 5 minutos de incubação. O processo inicia-se pela clivagem do fragmento N-terminal, região pró-peptídica contendo o motivo ERFNIN, envolvido na inibição da enzima madura. O meio ácido proporciona a auto-ativação, pois enfraquece as ligações de interação entre a região inibitória com o sítio catalítico. Neste contexto, BROMME et al., (2004), recomenda que o tempo de incubação deve ser o necessário para determinar o nível da atividade proteolítica, uma vez que a autocatálise, assim como a ativação, leva a uma parcial degradação da enzima recombinante. Assim, o tempo ótimo de encubação deve ser determinado e este deve ser o ponto de equilíbrio entre a ativação máxima e a degradação mínima da enzima. A ativação ou o processamento auto-catalítico pode ser parado, neste caso, aumentando o pH.

A pró-enzima de trabalho foi incubada em tampão 100 mM de acetato de sódio, 5 mM de DTT e pH 4.5 no período de 5 a 35 minutos. O processo de ativação (clivagem e digestão do pró-domínio N-terminal) foi analisado em SDS-PAGE Figura 4.14-A. Nos primeiros 5 minutos já foi possível visualizar a atividade proteolítica da enzima e após os 20 minutos, observa-se a sua atividade máxima. Aos 30 minutos é possível notar uma banda aparente em torno de 28 kDa, provável enzima madura sem a região inibitória (pró-domínio). As catepsinas L, depois de removida a região que as inibem, costumam ser um monômeros com massa molecular de 20 a 35 kDa (TURK et al., 1998 ; SHENAI et al., 2000; FONSECA et al., 2012). A região pró-peptídica foi totalmente processada após os 35 minutos e pôde ser monitorada no gel SDS. O processo de ativação da enzima a partir da auto-proteólise da região pró-peptídica N-terminal foi acompanhado por *immunoblotting*. A Figura 4.15 ilustra o período de incubação.

FIGURA 4.13 - Análise em SDS PAGE da proteína incubada em pH ácido, variando tempo de incubação e reconhecimento do anticorpo anti-His.

A: MM: Marcador de massa molecular (Thermo Unstained Protein MW Marker - Thermo Scientific; 1: pellet; 2: pró-enzima com 52 kDa, indicado por seta de cor preta; 5 min a 35 min: período de incubação da pró-enzima. A seta vermelha tracejada, sugere que seja o monômero da catepsina L com massa aparente de 28 kDa. **B** - Western blot usando o anti-His - 1: pellet; 2: pró-enzima com 52 kDa, reconhecida pelo anticorpo anti-His indicado por seta de cor preta; 5 min a 35 min: período de incubação da pró-enzima. A seta tracejada indica uma banda em torno de 17 kDa que sugere ser a região N-terminal clivada contendo partes dos aminoácidos do pró-peptídeo e as tags trx e his, marcado pelo anticorpo.

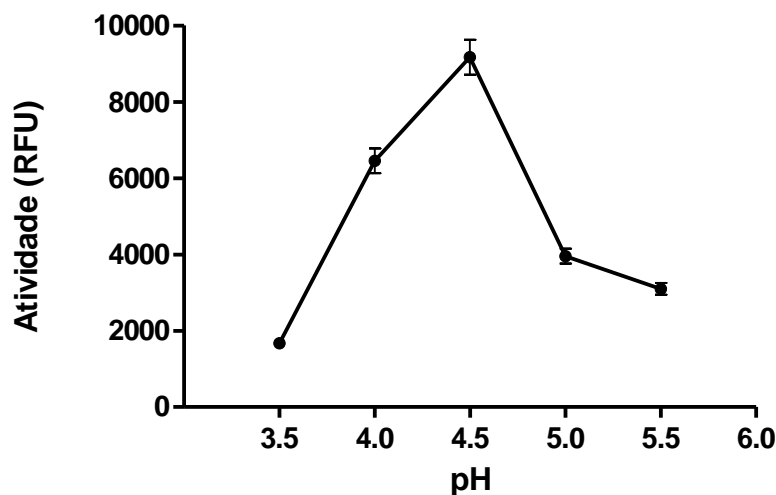


Fonte: Própria

4.5 Caracterização cinética da enzima AsCathL

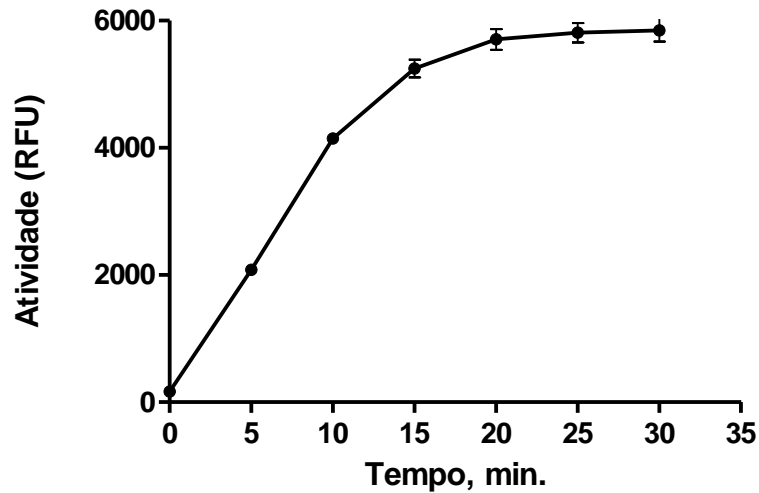
A hidrólise do substrato (Z-FR-MCA) pela AsCathL foi analisada em tampão 100 mM acetato de sódio, numa faixa de pH (3,5 – 5,5). A enzima foi ativada na presença de 5 mM de DTT, e após cinco minutos de incubação a atividade da enzima foi monitorada em função do pH, Figura 4.16. A atividade enzimática é dado pelo aumento da fluorescência (hidrólise do substrato Z-FR-MCA) detectada por espectrofluorímetro em unidade relativa de fluorescência (RFU). A enzima apresentou maior atividade proteolítica em pH 4,5. Este valor é similar ao encontrado para catepsinas L de insetos (PYATI et al., 2009; WATANABE et al., 2004) e a preferência da enzima por pH ácido está de acordo com o observado para catepsinas L de insetos na literatura (PHILIP et al., 2007; FONSECA et al., 2012). A Figura 4.17 exhibe o comportamento enzimático durante o período de incubação, nota-se que após 20 minutos, a atividade enzimática sobre o substrato é máxima, mantendo-se até os 30 minutos, tempo final estabelecido para o monitoramento da fluorescência da molécula 7 amino 4 metilculmarina (MCA) livre dado em unidade de fluorescência relativa (RFU).

FIGURA 4.14- Atividade da AsCathL dependente do pH usando o Z-Phe-Arg-AMC como substrato



Fonte: Própria

4.15 - Efeito do tempo de pré-incubação em pH 4,5 na atividade de AsCathL recombinante. Atividade hidrolítica da AsCathL contra o substrato fluorgênico no intervalo de 30 minutos



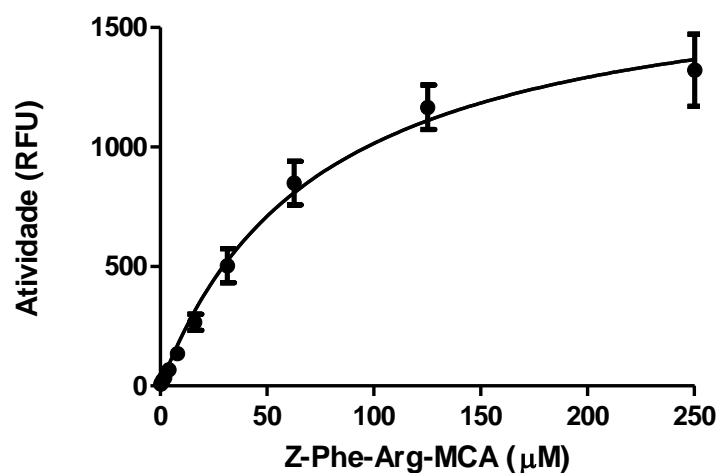
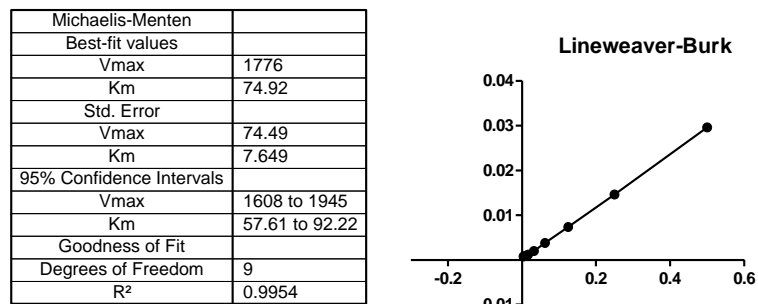
Fonte: Própria

A curva de hidrólise do substrato Z-Phe-Arg-AMC pela enzima AsCathL, a partir do qual o valor de K_m foi calculado, está ilustrada na Figura 4.18. O valor obtido para a constante de Michaelis–Menten, K_m foi de $74,92 \pm 7,6 \mu\text{M}$. Os valores encontrados na literatura para catepsinas recombinantes de insetos quando o substrato fluorgênico Z-FR-MCA foi usado é variável. O valor do K_m de $50 \mu\text{M}$ para a catepsina L digestiva do inseto *Tenebrio molitor* (CRISTOFOLETTI et al., 2005) e do K_m de $231 \mu\text{M}$ para o inseto *Bombyx mori* (WATANABE et al., 2004). Esses valores de K_m por exemplo, mudam quando a catepsina do *T. molitor* hidrolisa o substrato Z-Arg-Arg-MCA, obtendo K_m de $140 \mu\text{M}$, enquanto para a catepsina do inseto *B. mori* o valor de K_m obtido é de $40 \mu\text{M}$ quando o substrato Z-Leucina-Arginina-MCA foi utilizado. Estes valores mostram a especificidade do substrato frente à catepsina L.

É importante lembrar a importância da especificidade dos substratos em relação aos sítios ativos das peptidases. Estes sítios são divididos em subsítios localizados na região do substrato onde ocorre a hidrólise peptídica e são identificados como S1, S2, S3...Sn na direção N-terminal e S1', S2', S3', ..., Sn' na direção C-terminal da enzima. As posições correspondentes no substrato são: P1, P2, P3 ...Pn e P1', P2', P3' ...Pn' (SCHECHTER e BERGER 1967). A eficiência com que a enzima purificada funciona pode depender do tamanho do substrato (CRISTOFOLETTI et al., 2005). O substrato que não ocupa o subsítio S2 da enzima é pobremente hidrolisado. A diferença é causada pelo alto valor de K_m e o baixo

valor de k_{cat} . O substrato Z-Phe-Arg-MCA é o mais usual em catepsinas L, porém não se pode fazer dessa preferência uma regra, uma vez que algumas catepsinas de insetos preferem o resíduo de leucina no sítio P2 ao resíduo da fenilalanina. Apesar de ambos os aminoácidos terem cadeias laterais hidrofóbicas, o que os faz diferentes é a presença de um anel aromático na fenilalanina o que dá a esse aminoácido um tamanho maior em relação à leucina (FONSECA et al., 2012). Neste trabalho, apenas o substrato Z-FR-MCA foi utilizado para acompanhar a atividade enzimática da catepsina AsCathL.

FIGURA 4.16- Curva para determinação da constante de Michaelis-Menten (K_m) do substrato Z-Phe-Arg-MCA frente à catepsina L (AsCathL)



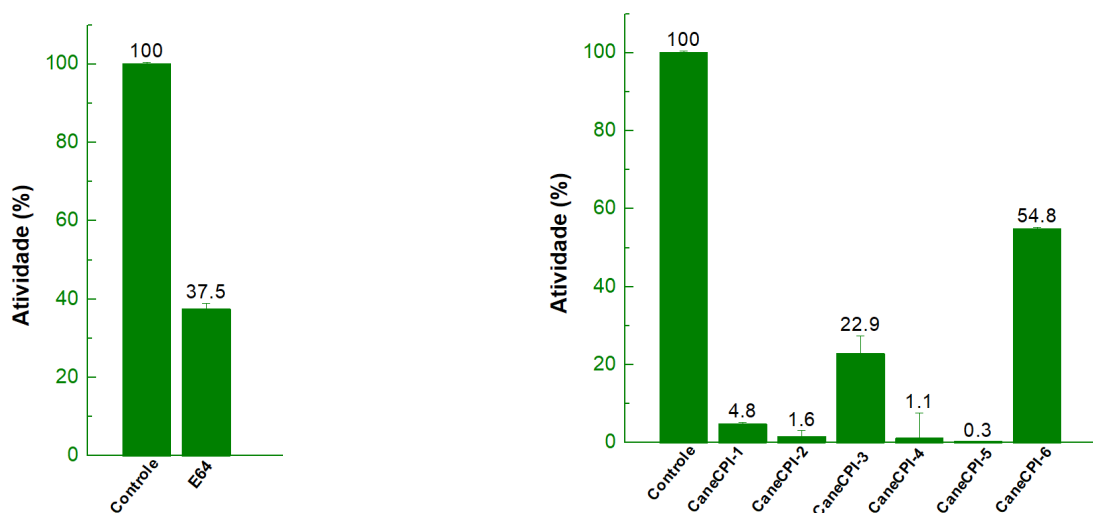
Fonte: Própria

4.6 Inibição da atividade da enzimática da AsCathL

Os testes de inibição foram feitos com o inibidor epóxido E-64 e as cistatinas recombinantes de cana-de-açúcar (canacistatinas), CaneCPI-1 (13 kDa), CaneCPI-2 (12.2 kDa), CaneCPI-3 (24 kDa), CaneCPI-4 (10.8 kDa), CaneCPI-5 (14.5 kDa), CaneCPI-6 (13 kDa), proteínas recombinantes expressas e caracterizadas pelo laboratório do Prof. Dr. Fávio Henrique da Silva, do Departamento de Genética e Evolução da UFSCar. Como mencionado anteriormente as CaneCPI-1, 2, 3 e 4 apresentaram atividade inibitória contra a catepsina L digestiva, rSl-CathL, do inseto praga da cana-de-açúcar, *Sphenophorus levis* (FONSECA et al., 2012), e, mais recentemente, a validação da ação inibitória da CaneCPI-4 contra a cisteíno peptidase do protozoário *Plasmodium falciparum* (MELO et al., 2018).

O resultado da triagem feita com esses inibidores mostra uma porcentagem de ação inibitória satisfatória sobre a atividade da enzima AsCathL, o qual pode ser visualizado na Figura 4.19.

FIGURA 4.17 - Avaliação da interferência do epóxido E-64 e das cistatinas recombinantes de cana-de-açúcar (CaneCPI-1, CaneCPI-2, CaneCPI-3, CaneCPI-4, CaneCPI-5, CaneCPI-6) na atividade da catepsina L (AsCathL)



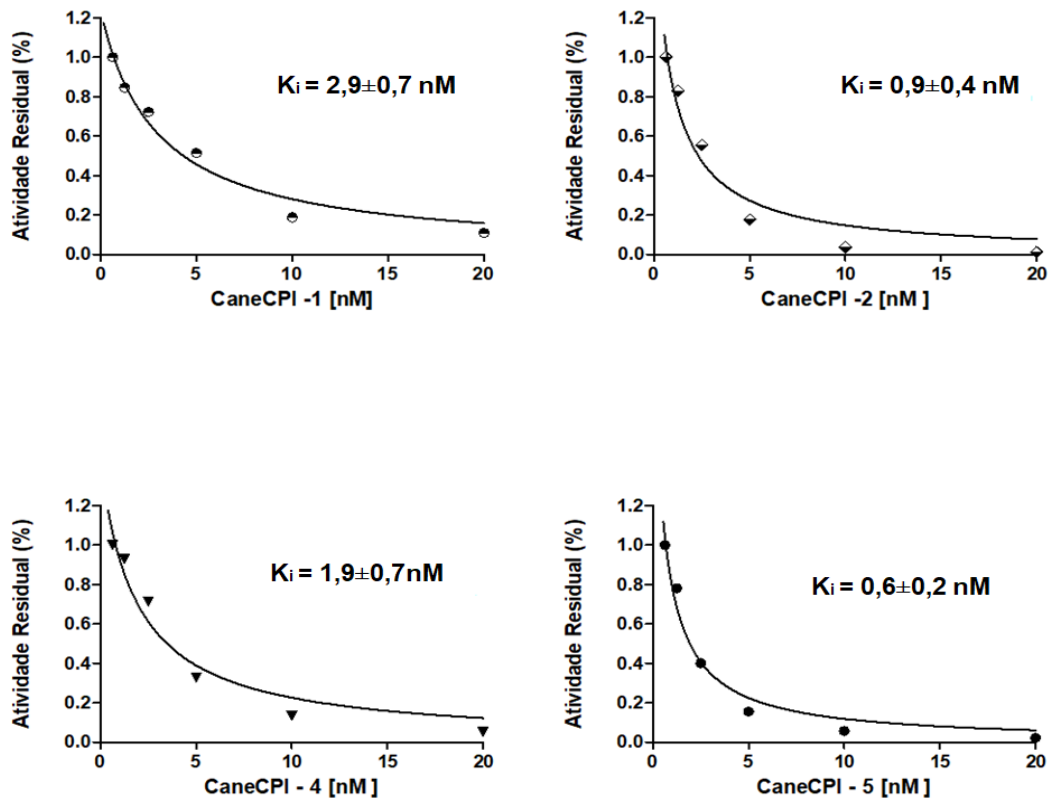
Fonte: Própria

Os inibidores cistatínicos, CaneCPI-1, CaneCPI-2, CaneCPI-4, CaneCPI-5 mostraram mais de 90% de inibição. O inibidor E-64 mostrou uma inibição menor quando comparado com alguns inibidores recombinantes, isto pode

ser explicado pela diferença no mecanismo de ação. O epóxido E-64 é o inibidor clássico de cisteíno peptidase e espera-se que a E-64 tenha menos especificidade para cisteíno peptidase digestiva quando comparado com as proteínas de defesa da cana-de-açúcar, as canacistatinas, que provavelmente evoluíram contra a classe das peptidases digestivas (FONSECA et al. , 2012). A concentração molar dos inibidores (canacistatinas) utilizados na reação final foi de 20 nM, enquanto do E-64 foi de 250 nM para uma taxa de inibição final de 62.5%,

As canacistatinas CaneCPI-3 e CaneCPI-6 tiveram uma porcentagem de inibição menor quando comparada com as demais, e a ação dessas proteínas sobre a AsCathL pode ser explicada pela diferença dos aminoácidos ao longo da sequência dessas proteínas e pelo tamanho da sequência primária, podendo diferir o mecanismo de ação de uma pra outra. O domínio dessas proteínas segundo alinhamento feito pelo grupo de pesquisas do Prof. Dr. Flávio mostrou ser bem conservado, mas com diferenças ao longo das sequências de aminoácidos. A CaneCPI-6, por exemplo, tem sua região C-terminal maior, quando comparada às canacistatinas CaneCPI- 1, 2, 4 e 5, porém essa extensão é menor quando comparada à CaneCPI-3. Dentre essas cistatinas, as que possuem maior similaridade em sua estrutura primária são as proteínas CaneCPI-4 e CaneCPI-5 (MIGUEL, 2014). A Figura 4.20 exibe os gráficos da inibição e o valor da constante de inibição (K_i) para os inibidores que apresentaram 90% de inibição sobre a AsCathL. Segundo COPELAND (2005), para um inibidor ser potente, sua afinidade pela enzima alvo deve ser alta, tendo valor de K_i menor ou igual à concentração total da enzima na reação. O valor da concentração da enzima utilizados nos ensaios foi de 9 nM, e os valores de K_i (Cane1 = $2,95 \pm 0,7$ nM; Cane2 = $0,92 \pm 0,4$ nM; Cane4 = $1,88 \pm 0,7$ nM e Cane5 = $0,6 \pm 0,2$ nM), exibidos sugerem que os inibidores analisados foram eficientes na inibição da catepsina e inibem de maneira igual a enzima, uma vez que estão na mesma ordem de grandeza.

FIGURA 4.18 - Efeito das canacistatinas na atividade da AsCathL



Fonte: Própria

4.7 Identificação dos aminoácidos da catepsina por LC/MS

Os peptídeos trípticos da proteína purificada (AsCathL) depois de clivados sob a ação da tripsina foram analisados por LC/MS. Os espectros foram pesquisados contra o banco de dados NCBI^{nr} utilizando o software Mascot e contra a sequência proteica de interesse.

Na análise foi possível identificar 14 peptídeos (em vermelho na sequência predita de AsCathL) na Figura 4.21 com uma cobertura de 60% sobre a sequência da proteína purificada. O peso molecular teórico calculado para a AsCathL foi de 53.566 Da, valor próximo da banda visualizada em SDS-PAGE em torno de 52 kDa.

FIGURA 4.19 - sequência de nucleotídeos da construção da AsCathL. Em vermelho, os peptídeos trípticos identificados por espectrometria de massas.

Protein View: sp|A00001|CatL_construction

sp|A00001|CatL_construction

Database: Labiommi_prot
 Score: 32811
 Nominal mass (M_r): 53566
 Calculated pI: 6.24

Sequence similarity is available as [an NCBI BLAST search of sp|A00001|CatL_construction against nr.](#)

Search parameters

MS data file: UFSCar_Amostra_K.mgf
 Enzyme: Trypsin: cuts C-term side of KR unless next residue is P.
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)

Protein sequence coverage: 60%

Matched peptides shown in **bold red**.

```

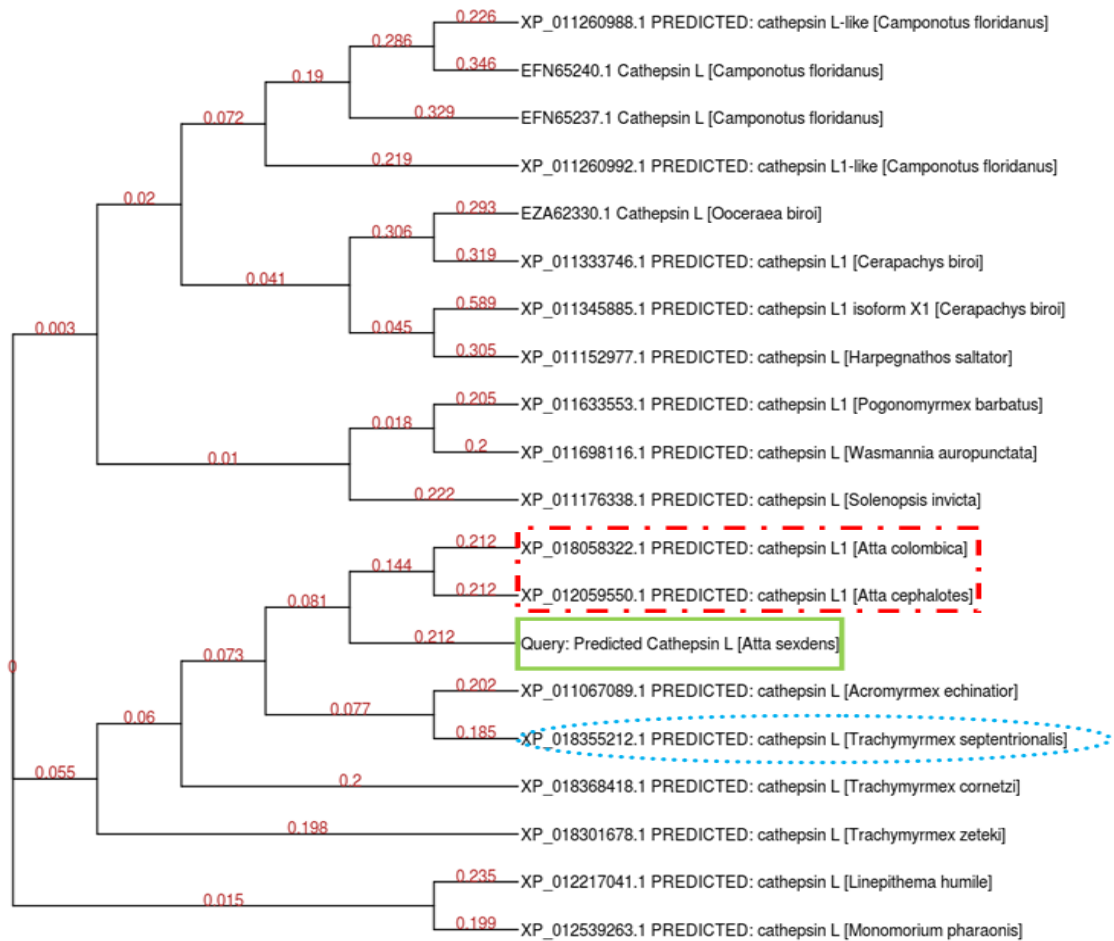
1  MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY
51  QGKLTIVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL
101 KEFLDANLAL AGSGSGHMH HHHSSGLVP RSGGMKETAA AKFERQHMS
151 PDLGTDDDDK AISFNKILDV EWFSFKTNHK KIYKSSVEEG LRMKIFLNNK
201 HKIIEHNRKY KLNEVPYKLG MNKYSDMLHH EFVNTLNGFN KSVKAQKQFM
251 GATFISPANV ELPKEVDWRK HGAVTEVKDQ GHCGSCWAFS TTGSLEGQHF
301 RQTGILVSL S EQNLIDCSGK YGNEGCNGGL MDNAFKYVRD NKGLDTEKSY
351 PYEAEKDKCR YNPRNSGATD VGFVDIPRGN EEKLKAAVAT IGFVSVDAID
401 SHESFQFYSE GVFYDFECNT ENLDHGVLIV GYGTDSKTGH DYWLKNSWG
451 ETWGEKGYIK MARNKDNHCG IASSASYP
  
```

Fonte: *software* Mascot

4.8 Análise dos dendrogramas

A árvore filogenética construída a partir do alinhamento realizado no ClustalW mostra que a catepsina L da formiga *Atta cephalotes* (XP_012059550) e *Atta colombica* (XP_018058322) são próximas em identidade com a catepsina L da *Atta.sexdens*, ambas pertencentes ao mesmo gênero. A Figura 4.21 ilustra o dendrograma gerado a partir do resultado do alinhamento múltiplo entre os aminoácidos de preditas catepsinas L de doze gêneros de formigas e a catepsina de trabalho. Na caixa tracejada em vermelho, estão as catepsinas (*A.colombica* e *A.cephalotes*). O quadrado em verde identifica a sequência estudada, e a caixa tracejada em azul, a predita catepsina L de *T.septentrionalis*.

Figura 4.20- Análise filogenética entre as catepsinas L das formigas depositadas no banco de dados NCBI e a predita catepsina da formiga *A.sexdens*



Fonte: programa R - <https://www.r-project.org/>

O resultado é condizente com a análise feita no NCBI utilizando o BLASTp, onde verificou-se que os aminoácidos da *A.sexdens* possuem 94% de identidade com a *Atta cephalotes*, 93% com a formiga *Atta colombica* e 91% com a formiga *Acromyrmex echinator*. O resultado do alinhamento mostrou que algumas sequências das catepsinas das formigas como a do gênero *Camponotus* (82%), possuem alta similaridade quando comparadas com a catepsina L da *A. sexdens*, porém a relação evolutiva está distante quando relacionadas com gênero *Atta*. A formiga *Trachymyrmex septentrionalis* possui 93% de identidade e está mais próxima do gênero *Atta*, Figura 4.22. Dos doze gêneros de formigas estudadas, oito delas já possuem seus genomas depositados no NCBI (*Atta cephalotes*, *Acromyrmex echinator*, *Solenopsis invicta*, *Harpegnathos saltator*, *Pogonomyrmex barbatus*, *Cerapachys biroi*, *Linepithema humile* e *Monomorium pharaonis*).

Figura 4.21- Porcentagem de identidade entre as sequências estudadas

Description	Max score	Total score	Query cover	E value	Ident
XP_012059550.1 [Atta cephalotes]	640	640	99%	0.0	94%
XP_018355212.1 [Trachymyrmex septentrionalis]	632	632	99%	0.0	93%
XP_018058322.1 [Atta colombica]	632	632	99%	0.0	93%
XP_018368418.1 [Trachymyrmex cornetzi]	625	625	99%	0.0	91%
XP_018301678.1 [Trachymyrmex zeteki]	621	621	99%	0.0	91%
XP_011067089.1 [Acromyrmex echinator]	602	602	99%	0.0	91%
XP_012539263.1 [Monomorium pharaonis]	576	576	100%	0.0	83%
XP_011698116.1 [Wasmannia auropunctata]	574	574	99%	0.0	83%
XP_011633553.1 [Pogonomyrmex barbatus]	571	571	99%	0.0	83%
XP_011260992.1 [Camponotus floridanus]	570	570	99%	0.0	82%
XP_011176338.1 : [Solenopsis invicta]	566	566	99%	0.0	82%
XP_012217041 [Linepithema humile]	565	565	99%	0.0	80%
EFN65237.1 [Camponotus floridanus]	541	541	95%	0.0	82%
XP_011152977.1 [Harpegnathos saltator]	521	521	99%	0.0	75%
XP_011333746.1 [Cerapachys biroi]	511	511	99%	0.0	74%
EZA62330.1 [Ooceraea biroi]	493	493	94%	0.0	76%
EFN65240.1 [Camponotus floridanus]	426	426	84%	2E-154	75%
XP_011260988.1 [Camponotus floridanus]	423	423	83%	3E-154	76%
XP_011345885.1 [Cerapachys biroi]	221	221	99%	2E-73	37%

Fonte: NCBI <https://www.ncbi.nlm.nih.gov>

Com o objetivo de detectar similaridade nas sequências analisadas e classificar quanto às funções biológicas das catepsinas nos organismos dos insetos, a segunda árvore filogenética foi construída. Um alinhamento entre as sequências das catepsinas de insetos de diferentes ordens e a catepsina L de *A. sexdens* foi feito utilizando o Clustalw e a árvore filogenética construída a partir desse alinhamento.

A análise filogenética entre a catepsinas L de *A. sexdens* e as catepsinas de insetos é exibida na Figura 4.24. A análise envolveu o alinhamento dos aminoácidos das catepsinas de 13 insetos, distribuídas em 4 ordens (Coleoptera, Diptera, Lepidoptera e Hymenoptera). Na caixa tracejada em azul, está a identificação dos insetos que possuem catepsinas L com identidade alta comparada com a catepsina L de *A. sexdens*. O quadrado em verde identifica a catepsina L de trabalho. De acordo análise feita no NCBI BLASTp, a predita catepsina L em *Apis mellifera* (XP_625135) possui 74% de identidade com a catepsina L da *A. sexdens*, e a *Nasonia vitripennis* (XP_001602523) possui 72%, Figura 4.23, ambas as sequências dos insetos pertencem à ordem *Hymenoptera*, e ambas sem caracterização funcional. As catepsinas da ordem mais próxima da *Hymenoptera*, são as da ordem Lepidoptera, conhecidas por serem atuantes na degradação da cutícula do inseto durante a muda (LIU et al., 2006) e na digestão da fibroína, proteína da seda (WATANABE 2004). Seguindo para grupos mais distantes, a

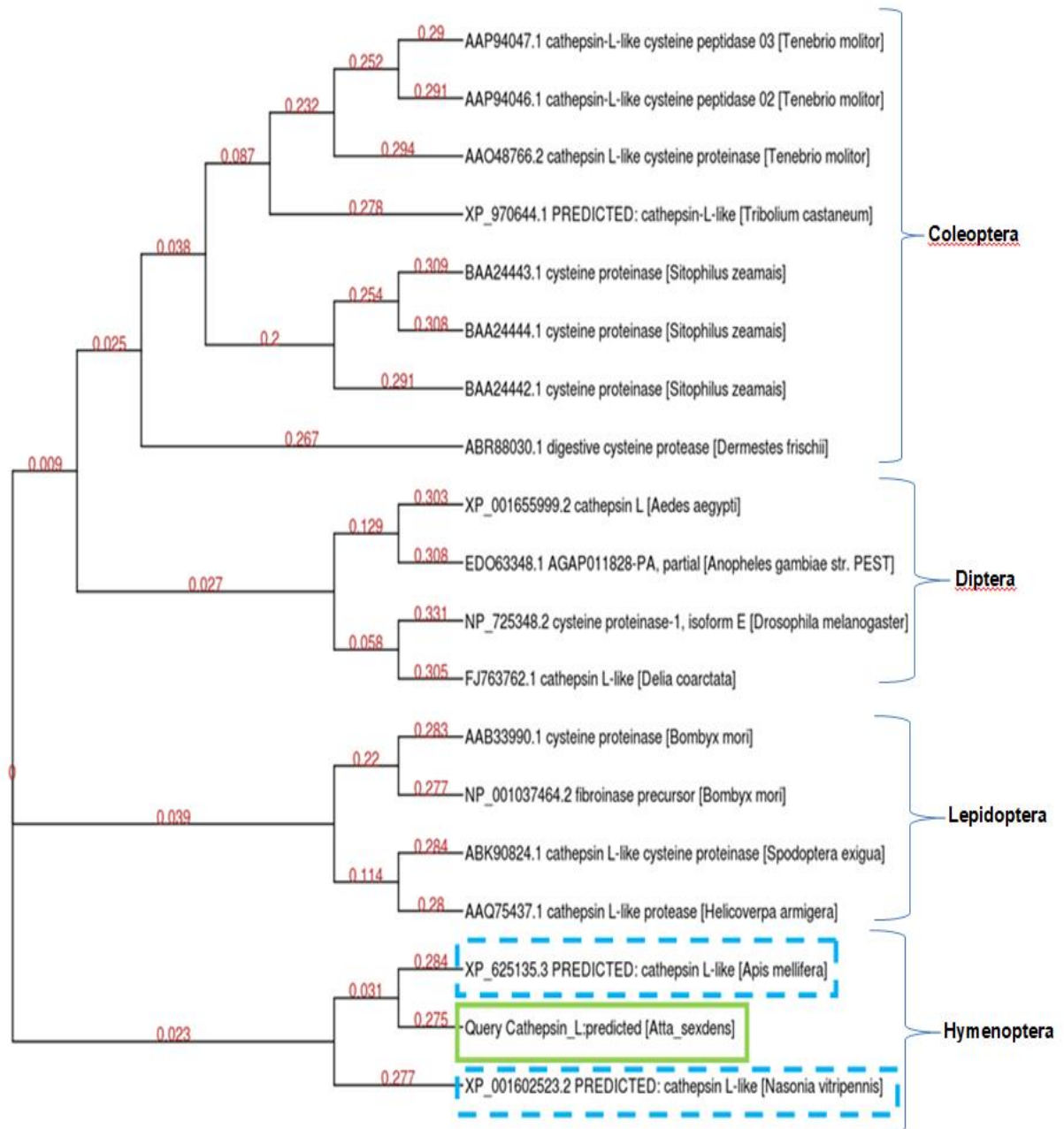
Coleoptera possui catepsinas caracterizadas como digestiva, mas também como lisossomais (CRISTOFOLETTI et al., 2005), e para a ordem Diptera, as catepsinas foram caracterizadas como lisossomais, no caso da mosca-da-fruta *D. melanogaster* a catepsina é expressa tanto na larva quanto no inseto adulto.

Figura 4.22- Porcentagem de identidade

Description	Max score	Total score	Query cover	E value	Ident
XP_625135.3 [<i>Apis mellifera</i>]	522	522	99%	0.0	74%
ABR88030.1 [<i>Dermestes frischii</i>]	484	484	99%	7E-177	69%
NP_001037464.2 [<i>Bombyx mori</i>]	479	479	99%	6E-175	69%
FJ763762.1 [<i>Delia coarctata</i>]	476	476	99%	6E-174	67%
XP_970644.1 [<i>Tribolium castaneum</i>]	476	476	99%	1E-173	68%
AAQ75437.1 [<i>Helicoverpa armigera</i>]	475	475	99%	3E-173	69%
EDO63348.1 [<i>Anopheles gambiae</i>]	472	490	99%	4E-172	65%
BAA24442.1 [<i>Sitophilus zeamais</i>]	470	470	99%	2E-171	68%
AAB33990.1 [<i>Bombyx mori</i>]	469	469	99%	1E-170	67%
XP_001655999.2 [<i>Aedes aegypti</i>]	468	468	99%	2E-170	66%
ABK90824.1 [<i>Spodoptera exigua</i>]	465	465	99%	2E-169	67%
AAP94046.1 [<i>Tenebrio molitor</i>]	454	454	99%	7E-165	63%
AAP94047.1 [<i>Tenebrio molitor</i>]	453	453	99%	1E-164	63%
AAO48766.2 [<i>Tenebrio molitor</i>]	452	452	99%	2E-164	63%
NP_725348.2 [<i>Drosophila melanogaster</i>]	450	450	94%	2E-163	67%
BAA24444.1 [<i>Sitophilus zeamais</i>]	442	442	97%	1E-160	66%
BAA24443.1 [<i>Sitophilus zeamais</i>]	441	441	97%	5E-160	65%
XP_001602523.2 [<i>Nasonia vitripennis</i>]	446	791	99%	2E-159	72%

Fonte: NCBI <https://www.ncbi.nlm.nih.gov>

Figura 4.23 - Análise filogenética entre as catepsinas L dos insetos depositadas no banco de dados NCBI e a predita catepsina da formiga *A.sexdens*



Fonte: programa R - <https://www.r-project.org/>

5 CONSIDERAÇÕES FINAIS E PERSPECTIVAS

O resultado da análise qualitativa do gene codificante para catepsina L na formiga cortadeira *A.sexdens* mostrou que ele é expresso nos estágios de desenvolvimento da formiga (larva, pupa, adulto), e está presente em várias partes da formiga operária (cabeça, tórax, gaster). A amplificação do gene a partir do DNA genômico extraído da cabeça da formiga nos sugere que a enzima estudada é do organismo da formiga e não do fungo *L. gongylophorus* com quem vive em simbiose.

A estrutura primária da proteína após sequenciamento foi identificada como uma cisteíno peptidase, catepsina L, pertencente à família C1, clã CA, que tem como principal representante a papaína. Além disso, a expressão recombinante da proteína AsCathL realizada em *E.coli* proporcionou uma expressão acentuada de proteína em corpos de inclusão, o que nos levou a “resgatar” essas proteínas através do *refolding*, técnica associada à purificação em coluna de afinidade, proporcionando um rendimento de 5 mg/L de proteína. A análise por espectrometria de massas, foi capaz de identificar os peptídeos da proteína purificada em solução com uma cobertura de 60% sobre a sequência de aminoácidos, indicando uma massa molecular teórica de 53 kDa, consistente com aquela visualizada em gel SDS PAGE em torno de 52 kDa, detectada por *immunoblotting*, utilizando anticorpo anti-his.

A ativação da proteína purificada permitiu a remoção do pró-peptídeo N- terminal da AsCathL sob condições ácidas e mostrou estar ativa hidrolisando o substrato fluorgênico Z-FR-MCA. Com a caracterização da enzima foi possível conhecer o seu comportamento proteolítico frente a diferentes pH assim como a velocidade com que essa enzima hidrolisa o substrato. Os ensaios de inibição da atividade enzimática realizados com os inibidores de cisteíno peptidases de cana-de-açúcar (6 canacistatinas) mostraram que estes são excelentes inibidores da catepsina L.

Os alinhamentos múltiplos mostram que a sequência da catepsina L da *A.sexdens* possui uma similaridade alta com as outras catepsinas de formigas do mesmo gênero. Dessa análise não foi conclusivo saber qual seria a função biológica da catepsina L no organismo de *A.sexdens*, pois as formigas que possuem genoma

completo depositado no banco de dados ainda não possuem catepsinas caracterizadas.

Em suma, consideramos que as etapas propostas inicialmente para o projeto foram realizadas com sucesso, proporcionou uma quantidade de informações científicas significativa sobre a cisteína peptidase em estudo. Este trabalho é precursor na caracterização da catepsina de formiga *A. sexdens*, levando a várias vertentes que podem ser tomadas para se seguir no estudo dessa enzima. Como as catepsinas estão envolvidas tanto no desenvolvimento como na sobrevivência dos insetos e estudos relatam a ação inseticida da catepsina sobre outros insetos, o conhecimento sobre a localização e a função biológica no organismo do inseto, e dos mecanismos envolvidos na atuação dessa enzima podem levar a estudos mais direcionados objetivando o controle do inseto.

O trabalho gerou colaborações e novos conhecimentos, como a colaboração com o Prof. Dr. Flávio, com os inibidores proteicos cedidos, foi possível fazer os ensaios de inibição da enzima, nos proporcionando novos e específicos conhecimentos, úteis para compreensão do mecanismo de ação e inibição das cisteína peptidase.

Com o objetivo de propor uma estrutura tridimensional ou, talvez, um *ensemble* estrutural para a Catepsina L, está sendo realizada a modelagem da estrutura da proteína por métodos *ab initio*. Nesse sentido, as principais opções para realizar a predição são as abordagens baseadas em conhecimento (*template-based modelling*) e *ab initio*. A validação da estrutura proposta será realizada ao fim do processo de modelagem a partir de dados experimentais fornecidos pela técnica de ligação cruzada acoplada a espectrometria de massas. Tendo em vista a necessidade de realização dos experimentos de ligação cruzada em valores de pH em torno de 4,5-5,0, faixa de atividade ótima da catepsina estudada. Este trabalho está sendo realizado em colaboração com o Dr. Hugo de Jesus.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ABE, K.; KONDO, H.; ARAI, S. "Purification and characterization of a rice cysteine proteinase inhibitor". *Agricultural and Biological Chemistry*, **51**: 2763-2768, 1987.

ABRAHAMSON, M.; ALVAREZ-FERNANDEZ, M.; NATHANSON, C. M. "Cystatins". *Biochem. Soc. Symp.*, 3: 179–199, 2003.

ALVES, L.C.; ALMEIDA, P.C.; FRANZONI, L.; JULIANO, L.; JULIANO, M.A. "Synthesis of N-alpha-protected aminoacyl 7-amino-4-methylcoumarin amide by phosphorus oxychloride and preparation of specific fluorogenic substrates for papain". *Peptide Res.*, 9: 92–96, 1996.

ARAÚJO, M. S.; RODRIGUES, C. A.; OLIVEIRA, M. A.; JESUS, F. G. "Controle biológico de formigas-cortadeiras: o caso da predação de fêmeas de *Atta* spp. por *Canthon virens*". *Revista de Agricultura Neotropical*, 2(3): 8–12, 2015.

BACCI J.R, M; SOLOMON, S.E.; MUELLER, U.G.; MARTINS, V.G.; CARVALHO, A. O.R.; VIEIRA, L.G.E.; SILVA-PINHATI, A.C.O. "Phylogeny of leafcutter ants in the genus *Atta fabricius (formicidae attini)* based on mitochondrial and nuclear DNA sequences". *Molecular Phylogenetics and Evolution*, **51**: 427– 437, 2009.

BARRET, A. J.; KEMBHAVI, A. A.; BROWN, M. A. "L-trans-Epoxy succinyl leucylamido (4-guanidínio) butane (E-64) and its analogues as inhibitor of cysteine proteinases including cathepsins B, H and L". *Biochemical Journal*, **201**: 189-198, 1982.

BARRET, A.J.; FRITZ, H.; GRUBB A.; ISEMURA, S.; JARVINEN, M.; KATUNUMA, N.; MACHLEIDT, W.; MULLERESTERL, W.; SASAKI, M; TURK, V. "Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin", *Biochem. J.* **236**: 312-312, 1986.

BACCI Jr., M.; ANVERSA, M.M.; PAGNOCCA, F.C. "Cellulose degradation by *Leucocoprinus gongylophorus*, the fungus cultured by the leaf-cutting ant *Atta sexdens rubropilosa*". *Braz. J. Med. Biol. Res.*, **28**:79–82, 1995.

BANEYX, F. "Recombinant protein expression in *Escherichia coli*". *Curr. Opin. Biotechnol*, **10**: 411 - 421, 1999.

BANEYX, F.; MUJACIC, M.; "Recombinant protein folding and misfolding in *Escherichia Coli*". *Nat Biotechnol*, 22(11): 1399-408, 2004.

BENCHABANE, M.; SCLUTER, U.; VORSTER, J.; GOULET, MR.; MICHAUD, D. "Plant cystatins". *Biochimie.* **92**: 1657-1666, 2010.

BENCKISER, G. "Ants and sustainable agriculture. A review. *Agronomy for sustainable development* **30**:191-199, 2010.

BERTENSHAW, G. P.; NORCUM, M. T.; BOND, J. S. "Structure of homo - and hetero-oligomeric meprin metalloproteases". *J. Biol. Chem.*, 278(4): 2522–2532, 2003.

BETON, D.; GUZZO, C.R.; RIBEIRO, A.F.; FARAH, C.S.; TERRA, W.R. "The 3d structure and function of digestive cathepsin L -like proteinases of *Tenebrio molitor* larval midgut". *Insect Biochemistry and Molecular Biology*, **42**: 655-664, 2012.

BIETH, J. G. "Theoretical and practical aspects of proteinase inhibition kinetics". *Methods Enzymol*, **248**: 59-84, 1995.

BODENHOFER, U.; BONATESTA, E.; HOREJŠ-KAINRATH, C.; HOCHREITER, S. "msa: an R package for multiple sequence alignment". *Bioinformatics*, 31(24): 3997–3999, 2015.

BOULOGNE, I.; OZIER-LAFONTAINE, H.; LORANGER-MERCIRIS, G. "Leaf-cutting ants, biology and control". *Sustainable Agriculture Reviews Agriculture Reviews*, 13. 2014.

BOULOGNE, I.; PETIT, P.; OZIER-LAFONTAINE, H.; DESFONTAINES, L.; LORANGER- MERCIRIS, G. "Insecticidal and antifungal chemicals produced by plants: a review". *Environ. Chem. Lett.*, **10**: 325–347, 2012.

BOYD, N.D.; MARTIN, M.M. " Fecal proteinases of the fungus growing ant *Atta texana*: their fungal origin and ecological significance". *Journal of Insect Physiology*, **21**: 1815–1820, 1975.

BRADFORD, M.M. "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding". *Analytical Biochemistry*, 72(1-2):248-254, 1976.

BROMME, D.; NALLASETH, F.S.; TURK, B. "Production and activation of recombinant *papain-like* cysteine proteases". *Methods*, **32**:199–206, 2004.

BLUM, G.; VON DEGENFELD, G.; MERCHANT, M. J.; BLAU, H. M.; BOGYO M. "Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes". *Nat. Chem. Biol.*, **3**: 668–77, 2007.

CAMARGO, R.S.; FORTI, L.C.; LOPES, J.F.S.; NAGAMOTO, N.S. "Studies on leaf-cutting ants, *Acromyrmex* spp. (Formicidae, *Attini*): behavior, reproduction and control". *Recent research developments in entomology*, **5**:1 - 21, 2006.

CANTARELLI, E.B.; COSTA, E.C.; PEZZUTTI, R.; OLIVEIRA, L.S. "Quantificação das perdas no desenvolvimento de *Pinus Taeda* após o ataque de formigas cortadeiras". *Ciência Florestal*, **18** (1): 39-45, 2008.

CANTÚ, M.D.; CARRILHO, E. "Sequenciamento de peptídeos usando espectrometria de massas: um guia prático". *Quim. Nova.* , 31(3): 669-675, 2008.

COSTA, L.G. "Current issues in organophosphate toxicology". *Clinica Chimica Acta*, **366**:1-13, 2006.

COULOMBE, R.; GROCHULSKI, P.; SIVARAMAN, J.; MENARD, R.; MORT, J.S.; CYGLER, M. "Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment". *Embo J.*, **15**:5492-5503, 1996.

COPELAND, R. A. "enzymes: a practical introduction to structure, mechanism, and data analysis". New York, Wiley, 2000.

CURRIE, C.R.; STUART, A.E. "Weeding and grooming of pathogens in agriculture by ants". *Proceeding of The Royal Society of London*, **268**: 1033 -1039, 2001.

CRISTOFOLETTI, P.T.; RIBEIRO, A.F.; TERRA, W.R."The cathepsin L-like proteinases from the midgut of *Tenebrio molitor* larvae: sequence, properties, immunocytochemical localization and function". *Insect Biochem. Mol. Biol.*, **35**: 883-901, 2005.

CLARK, E. "Protein refolding for industrial process". *Curr Opin Biotechnol*, **12**(2):202-7, 2001.

SCHNEIDER, V.K; SOARES-COSTA, A.; CHAKRAVARTHI, M; RIBEIRO, C; MOUTINHO, S.C.; FALCO, M.C.; HENRIQUE-SILVA, F. "Transgenic sugarcane overexpressing CaneCPI-1 negatively affects the growth and development of the sugarcane weevil *Sphenophorus levis*". *Plant Cell Reports*, **36**: 193-201, 2017.

CHO, WL; TSAO, SM; HAYS, AR; WALTER, R; CHEN, JS; SNIGIREVSKAYA, ES; RAIKHEL, S. "Mosquito cathepsin B – like protease involved in embryonic degradation of vitellin is produced as a latent extra ovarian precursor". *J. Biol. Chem.*, **274**:13311–13321, 1999.

CHOMCZYNSKI, P. & SACCHI, N. "Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction". *Analytical Biochemistry*, **162**(1): 156-159, 1987.

CHU, M.H.; LIU, K.L.; WU, H.Y.; YEH, K.W.; CHENG, Y.S. "Crystal structure of tarocystatin-papain complex: implications for the inhibition property of group-2 phytocystatins". *Planta*, **234**: 243-254, 2011.

CUOZZO, J.W.; TAO, K.; CYGLER, M.; MORT, J.S.; SAHAGIAN, G.G. "Lysine-based structure responsible for selective mannose phosphorylation of cathepsin D and cathepsin L defines a common structural motif for lysosomal enzyme targeting". *The Journal of Biological Chemistry*, **273**: 21067-21076, 1998.

DAMASCENO, T.C.; DIAS, R.O.; OLIVEIRA, J.R.; SALINAS, R.K.; JULIANO, M. A.; FERREIRA, C.; TERRA, W.T. "Active subsite properties, subsite residues and targeting to lysosomes or midgut lumen of cathepsins L from the beetle *Tenebrio molitor*". *Insect Biochemistry and Molecular Biology*, **89**: 17-30, 2017.

DUBIN, G. "Proteinaceous cysteine protease inhibitors". *Cell. Mol. Life Sci.*, **62**: 653–669, 2005.

DUFOR, E. "Sequence homologies, hydrophobic profiles and secondary structures of cathepsin S B; H and L: comparison with papain and actinidin". *Biochimie*, **70**: 1335-1342, 2001.

DURONIO, C. "Production of a recombinant vector to enable the study of thioredoxin function as a bound or detached solubilizer of proteinase inhibitor in a bacterial protein overexpression system". *J. Exp. Microbiol. Immunol.*, **16**:79-84, 2012.

DRAG, M.; SAIVESEN G.S. "Emerging principles in protease-based drug discovery". *Nat Rev Drug Discov*, **9**: 690–701, 2010.

FELLER, G.; LE BUSSY, O.; GERDAY, C. "Expression of psychrophilic genes in mesophilic hosts: assessment of the folding state of a recombinant α amylase". *Appl. Environ. Microbiol.*, **64**:1163–1165, 1998.

FERRER, M.; CHERNIKOVA, T.N.; YAKIMOVY, M.M.; GOLYSHIN, P.N.; TIMMIS, K.N. "Chaperonins govern growth of *Escherichia coli* at low temperatures". *Nat Biotechnol.*, 21(11):1266-7, 2003.

FONSECA, F. P.P.; SOARES-COSTA, A.; RIBEIRO, A. F.; ROSA, J. C.; TERRA, W. R.; HENRIQUE-SILVA, F. "Recombinant expression, localization and in vitro inhibition of midgut cysteine peptidase (SI-CathL) from sugarcane weevil, *Sphenophorus levis*". *Insect Biochemistry and Molecular Biology*, **42**: 58-69, 2012.

FOWLER, H.G. "Foraging trails of leaf-cutting ants". *J N Y Entomol Soc.*, **86**: 132–136, 1978.

GUAY, J.; FALGUEYRET, J.P.; DUCRETD, A.; PERCIVAL, M.D.; MANCINI, J.A. "Potency and selectivity of inhibition of cathepsin K, L and S by their respective propeptides". *Eur. J. Biochem.*, **267**: 6311-6318, 2000.

GIANOTTI, A.; SOMMER, C.A.; CARMONA, A.K.; HENRIQUE-SILVA, F. "Inhibitory effect of the sugarcane cystatin CaneCPI-4 on cathepsins B and L and human breast cancer cell invasion". *J. Biol. Chem.* **389**: 447-453, 2008.

GIANOTTI, A.; RIOS; W.M.; SOARES-COSTA, A.; NOGAROTO, V.; CARMONA, A.K.; OLIVA, M.L.V.; ANDRADE, S.S.; HENRIQUE-SILVA, F."Recombinant expression, purification, and functional analysis of two novel cystatins from sugarcane (*Saccharum officinarum*)". *Protein Express Purif.* **47**: 483 - 489, 2006.

GUNCAR, G.; PUNGERCIC, G.; KLEMENICIC, I., TURK, V.; TURK, D. "Crystal structure of MHC class II-associated p41 li fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S". *EMBO J*, **18**: 793-803, 1999.

GUPTA, R., JUNG, E. & BRUNAK, E. 'Prediction of N-glycosylation sites in human proteins" 2004 – Tipo referênci: trabalho não publicado, 2011.

HANAHAN, D. "Studies on transformation of *Escheriehia coli* with plasmid". *J. Moi. Biol*, **166**: 557-580, 1983.

- RAWLINGS, N. D.; BARRETT, A. J. "Evolution of proteins of the cystatin superfamily". *J. Mol. Evol.*, 30: 60–71, 1990.
- RIBEIRO, C.W.; SOARES-COSTA, A.; FALCO, M. C.; CHABREGAS, S.; ULIAN, E. C.; COTRIN, S. S.; CARMONA, A. K.; SANTANA, L.; OLIVA, M. L. V.; HENRIQUE-SILVA, F. "Production of a his-tagged Canecystatin in transgenic sugarcane and subsequent purification". *Biotechnology Progress*, 24: 1060-1066, 2008.
- HEGEDUS, D.; O'GRADY, M.; CHAMANKHAH, M.; BALDWIN, D.; GLEDDIE, S.; BRAUN, L.; ERLANDSON, M.. "Changes in cysteine protease activity and localization during midgut metamorphosis in the crucifer root maggot (*Delia radicum*)". *Insect. Biochem. Mol. Biol.* **32**: 1585–1596, 2002.
- HERNÁNDEZ, J.V.; JAFFÉ, K. "Dano econômico causado por populações de formiga *Atta laevigata* (F.Smith) em plantações de Pinus Carbaea Mor e elementos para o manejo da praga". *An. Soc. Entomol. Brasil*, 24(2):287-298, 1995.
- HERVEY, A.; ROGERSON, C.T.; LEONG, I. "Studies on fungi cultivated by ants". *Brittonia*, **29**: 226-236, 1997.
- HOMMA, K.; KURATA, S; NATORI, S. "Purification, characterization and cDNA cloning of procathepsin L from the culture medium of Nih-Sape-4, an embryonic cell line of *Sarcophaga peregrina* (Flesh Fly); and its involvement in the differentiation of imaginal discs". *J. Biol. Chem.*, **269**: 15258–15264, 1994.
- HUANG, R.; QUE, X.; HIRATA, K.; BRINEN, L. S.; LEE, J. H.; HANSELL, E.; ENGEL, J. ; SAJID, M.; REED, S. "The cathepsin L of *Toxoplasma gondii* (TgCPL) and its endogenous macromolecular inhibitor, toxostatin". *Mol Biochem Parasitol*, 164(1): 86-94, 2009.
- IPEF- Instituto de Pesquisas e Estudos Florestais - Documentos Técnicos IPEF, 4(4): 1-38, 2016.
- IKE, P.T.L.; MOREIRA, A.C.; ALMEIDA, F.G.; FERREIRA, D. ; BIROLI, W.G.; PORTO, A.L.M.; SOUZA, D.H.F. "Functional characterization of a yellow laccase from *Leucoagaricus gongylophorus*". *SpringerPlus*, **4**: 654, 2015.
- JULENIUS, K., MØLGAARD, A., GUPTA, R., BRUNAK, S. "Prediction, conservation analysis and structural characterization of mammalian mucin-type Oglycosylationsite". *Glycobiology*, **15**: 153, 2005.
- KARRER, M.; PEIFFER, S.L.; DITOMAS, M.E. "Two distinct gene subfamilies within the family of cysteine protease genes. *Proc Natl Sci USA*, **90**: 3063-3067, 1993.
- KONDO, H.; ABE, K. ; NISHIMURA, I.; WATANABE, H.; EMORI, Y.; ARAI, S. "Two distinct cystatin species in rice seeds with different specificities against cysteine proteinase". *Journal of Biological Chemistry*, **265**: 15832-15837, 1990.
- KONDO, H.; ABE, K.; EMORI, Y.; ARAI, S. "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 letters*, **278**: 87-90, 1991.

LAEMMLI, U.K. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4". *Nature*, **227**: 680-685, 1970.

LARSON, E.T.; PARUSSINI, F.; HUYNH, M.H.; GIEBEL, J.D.; KELLEY, A.M.; ZHANG, L.; BOGYO, M.; MERRITT, E.A.; CARRUTHERS, V.B. "Toxoplasma gondii cathepsin L is the primary target of the invasion inhibitory compound Morpholinurea-leucyl-homophenyl-vinyl sulfone phenyl". *J. Biol. Chem.*, **284**: 26839-26850, 2009.

LARKIN, M.A.; BLACKSHIELDS, G.; BROWN, N.P.; CHENNA, R.; MCGETTIGAN, P.A.; MCWILLIAM, H.; VALENTIN, F.; WALLACE, I.M.; WILM, A.; LOPEZ, R.; THOMPSON, J.D.; GIBSON, T.J.; HIGGINS, D.G. "Clustal W and Clustal X version 2.0". *Bioinformatics*, **23**(21): 2947–2948, 2007.

LECAILLE, F.; KALETA, J.; BROMME, D. "Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design". *Chem. Rev.*, **102**: 4459-88, 2002.

LI, H.; TANG, H.; SIVAKMAR, S.; PHILIP, J.; HARRISON, R.L.; GATEHOUSE, J.A.; BONNING, B.C. "Insecticidal activity of a basement membrane-degrading protease against *Heliothis virescens* (Fabricius) and *Acyrtosiphon pisum* (Harris)". *J. Insect Physiol.*, **54**: 777–789, 2008.

LIANG, C.; BROOKHART, G.; FENG, G. H.; REECK, G. R.; KRAMER, K. J. "Inhibition of digestive proteinase of stored grain Coleoptera by oryzacystatin, a cysteine proteinase inhibitor from rice seed. *FEBS Letters*, **278**(2): 139-142, 1991.

LIVRAMENTO, K.G.; FREITAS, N.C.; MÁXIMO, W.P.F.; ZANETTI, R.; PAIVA, L.V. "Gene Expression Profile Analysis is Directly Affected by the Selected Reference Gene: The Case of Leaf-Cutting *Atta Sexdens*". *Insects*, **9**(1): 18, 2018.

LIU, J.; SHI, G.P.; ZHANG, W.Q.; ZANG, G.R.; XU, W.H. "Cathepsin L function in insect moulting: molecular cloning and functional analysis in cotton bollworm, *Helicoverpa Armigera*". *Insect Mol Biol*, **15**: 823–834, 2006.

LÓPEZ-OTÍN, C.; BOND, J.S. "Proteases: multifunctional enzymes in life and disease". *Journal of Biological Chemistry*, **283**(45): 30433–30437, 2008.

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

MARGIS-PINHEIRO, M.; ZOLET, A.C.T.; LOSS, G.; PASQUALI, G.; MARGIS, R. "Molecular evolution and diversification of plant cysteine proteinase inhibitors: new insights after the poplar genome". *Mol. Phylogenet. Evol.*, **49**: 349-355, 2008.

MARTIN, M.M.; CARMAN, R.M.; MACCONNEL, J.G. "Nutrients derived from the fungus cultured by the fungus-growing ant *Atta colombica tonsipes*". *Annals of the Entomological Society of America*, **62**: 11–13, 1969.

MARTIN, M.M.; WEBER, N.A. "The cellulose utilizing capability of the fungus cultured by the *attini* ant *Atta colombica tonsipes*". *Annals of the Entomological Society of America*, **62**:1386–1387, 1969.

MATSUMOTO, I.; WATANABE, H.; ABE, K.; ARAI, S.; EMORI, Y. "A putative digestive cysteine proteinase from *Drosophila melanogaster* is predominantly expressed in the embryonic and larval midgut". *Eur. J. Biochem.*, **227**:582–587, 1995.

MATSUMOTO, I. ; EMORI, Y. ; ABE, K. , ARAI, S. "Characterization of a gene family encoding cysteine proteinases of *Sitophilus zeamais* (maize weevil), and analysis of the protein distribution in various tissues including alimentary tract and germ cells". *J. Biochem.*, **121**:464–476, 1997.

MELO, C.M. *Produção Recombinante e Caracterização de Duas Cistatinas de Cana-de-açúcar*. São Carlos, Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde - UFSCar, 2014. Dissertação, 42 p.

MOREIRA, A.C; FERREIRA, D.; ALMEIDA, F.G.; RODRIGUES-FILHO, E; FERNANDES, J.B.; SILVA, M.F.; VIEIRA, P.C.; PAGNOCCA, F.C.; SOUZA, D.H.F. "Molecular and kinetic characterization of two extracellular xylanases isolated from *Leucoagaricus gongylophorus*". *Appl Biochem Biotechnol.*, **173**(3):694-704, 2014.

MORRISON, J. F. "Kinetics of the reversible inhibition of enzyme catalyzed reactions of tight-binding inhibition". *Biochim. Biophys. Acta*, **185**:269-285, 1996.

NAGATA, K.; KUDO, N.; ABE, K.; ARAI, S.; TANOKURA, M. "Three-dimensional solution structure of oryzacystatin-I, a cysteine protease inhibitor of the rice, *Oryza sativa* L. japonica. *Biochemistry*, **39**: 14753-1476, 2000.

NICKELE, M.A.; REIS FILHO, W.; OLIVEIRA, E.B.; IEDE, E.T.; CALDATO, N.; STRAPASSON, P. "Leaf-cutting ant attack in initial Pine plantations and growth of defoliated plants". *Pesq. Agropec. Bras.*, **47**(7), 2012.

OLIVEIRA, A. S.; XAVIER-FILHO, J.; SALES, M. P. "Cysteine proteinase and cystatins". *Brazilian Archives of Biology and Technology*, **46**(1): 91-104, 2003.

OLIVA, M.L.; CARMONA, A. K.; ANDRADE, S.S.; COTRIN, S.S.; SOARES-COSTA, A.; HENRIQUE-SILVA, F. "Inhibitory selectivity of canecystatin: a recombinant cysteine peptidase inhibitor from sugarcane". *Biochem Biophys Res Commun.* **320**(4):1082–1086, 2004.

PARK, S.W.; PARK, S.T.; LEE, J.E.; KIM, Y.M." *Pseudonocardia* carboxydivorans sp. nov., a carbon monoxide-oxidizing actinomycete, and an emended description of the genus *Pseudonocardia*". *Int J Syst Evol Microbiol*, **58**(11):2475–2478, 2008.

PHILIP, J.M.; FITCHES, E.; HARRISON, R.L.; BONNING, B.; GATEHOUSE, J.A. "Characterization of functional and insecticidal properties of a recombinant cathepsin L-like proteinase from flesh fly (*Sarcophaga peregrina*), which plays a role in differentiation of imaginal discs". *Insect. Biochem. Mol. Biol.*, **37**: 589–600, 2007.

- POULSEN, M.; BOT, A. N. M.; BOOMSMA, J. J. "The effect of metapleural gland secretion on the growth of amutualistic bacterium on the cuticle of leaf-cutting ants". *Naturwissenschaften*, **90**:406–409, 2003.
- POWERS, J.C.; ASGIAN, J.L.; EKICI, O.D.; JAMES, K.E. "Irreversible inhibitors of serine; cysteine and threonine proteases". *Chem. Rev.*, **102**: 4639–4750, 2002.
- PYATI, P.S.; BELL, H.A.; FITCHES, E.; PRICE, D.R.G.; GATEHOUSE, A.M.R; GATEHOUSE, J.A. " Cathepsin L -like cysteine proteinase (Dccathl) from *Delia Coarctata* (wheat bulb fly): basis of insecticidal activity". *Insect Biochemistry and Molecular Biology*, **39**: 535–546, 2009.
- QUINLAN, R.J.; CHERRETT, J.M. "The role of fungus in the diet of the leaf-cutting ant *Atta cephalotes* (L.)". *Ecological Entomology*, **4**: 151-160, 1979.
- RAPPSILBER, J.; MANN, M.; ISHIHAMA, Y. " Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips". *Nat. Protoc.* **2**: 1896–190, 2007.
- RAWLINGS, N.D.; BARRETT, A.J.; BATEMAN, A. "Merops: the database of proteolytic enzymes, their substrates and inhibitors". *Nucleic Acids Research*, **40**: D343-D350, 2012.
- REIS, E.M.; MARGIS, R. "Sugarcane phytocystatins: identification, classification and expression pattern analysis". *Genet. Mol. Biol.*, **24**: 291–296, 2001.
- RENART, J.; REISER, J.; STARK, G.R. "Transfer of proteins from gels to diazobenzoyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure". *Proceedings of the National Academy of Sciences of the United States of America*, **76** (7): 3116, 1979.
- SANTIAGO, A.C.; KHAN, Z.N.; MIGUEL, V C.; GIRONDA, C.; SOARES-COSTA, V; PELÁ, V.T.; LEITE, V; EDWARDSON, J.M.; BUZALAF, M.A.R.; HENRIQUE-SILVA, F. "A new sugarcane cystatin strongly binds to dental enamel and reduces erosion". *Journal of Dental Research*, **96** (9): 1051–1057, 2017.
- SAMBROOK, J. & RUSSELL, D.W. *Molecular Cloning: "A laboratory manual.* 3^o ed. New York: Cold Spring Harbor Laboratory Press, 2001.
- SANMAN, L.E.; BOGYO, M. "Activity-based profiling of proteases". *Annu. Rev. Biochem*, **83**:249–73, 2014.
- SAITOU, N.; NEI, M. "The neighbor-joining method: a new method for reconstructing phylogenetic trees". *Molecular Biology and Evolution*, **4**(4): 406-425, 1987.
- SALVESEN, G. S.; NAGASE, H. "Inhibition of proteolytic enzymes. IN: BEYNON, R; BOND, J. S. *Proteolitic enzymes.* Oxford: Oxford University Press, 2001. p. 105-130.

SEVERINO, P.S. Busca de Produtos Naturais como Inibidores Específicos de Enzimas. São Carlos, Programa de Pós-Graduação em Química Orgânica – UFSCar, 2008. Tese de doutorado, 128 p.

SILVA, A.; BACCI, Jr. M. GOMES DE SIQUEIRA, C; BUENO, OC; PAGNOCCA, FC; HEBLING, MJA. “Survival of *Atta sexdens* workers on different food sources. Journal of insect Physiology, **49**:307–313, 2003.

SILVA-PINHATI, A.C.; BACCI, Jr. M.; HINKLE, M.L.; SOGIN, M.L.; PAGNOCCA, F.C.; MARTINS, V.G.; BUENO, O.C.; HEBLING, M.J.A. “ Low variation in ribosomal DNA and internal transcribed spacers of the symbiotic fungi of leaf-cutting ants (*Attini*: Formicidae)”. Braz. J. Med. Biol. Res., **37**: 1463–1472, 2004.

SILVA, A.; BACCI, Jr. M.; PAGNOCCA, F.C.; BUENO, O. C.; HEBLING, M.J.A. “ Starch metabolism in *Leucoagaricus gongylophorus*, the symbiotic fungus of leaf-cutting ants”. Microbiol Res., **161**:299-303, 2006.

SIQUEIRA, C.G.; BACCI Jr. M.; PAGNOCCA, F.C.; BUENO, O.C.; HEBLING, M.J.A. “Metabolism of plant polysaccharides by *Leucoagaricus gongylophorus*, the symbiotic fungus of leaf-cutting ant *Atta sexdens* L. Appl. Environ”. Microbiol, **64** (12): 4820–4822, 1988.

SOARES-COSTA, A.; BELTRAMINI, L.M. ; THIEMANN, O.H. ; HENRIQUE-SILVA, F. “A sugarcane cystatin: recombinant expression, purification, and antifungal activity, Biochem. Biophys. Res. Commun., **296**(5): 1194–1199, 2002.

SOARES-COSTA, A.; DIAS, A.B.; DELLAMANO, M.; DE PAULA, F.F.; CARMONA, A.K.; TERRA, W.R.; HENRIQUE-SILVA, F. “ Digestive physiology and characterization of digestive cathepsin L -like proteinase from the sugarcane weevil *Sphenophorus levis* ”. J insect Physiol., **57**(4): 462-468, 2011.

SONG, J.M.; AN, Y.J.; KANG, M.H.; LEE, Y.H.; CHA, S.S. “Cultivation at 6–10 degrees is an effective strategy to overcome the insolubility of recombinant proteins in *Escherichia coli*”. Protein Expres Purif., **82**: 297–301, 2012.

SOSNOWSKI, P.; TURK, D. “Caught in the act: the crystal structure of cleaved cathepsin L bound to the active site of cathepsin L”. FEBS Letters, **590**(8): 1253-61, 2016.

SUEN, G.; TEILING, C.; LI, L.; HOLT, C.; ABOUHEIF, E.; BORNBERG-BAUER, E.; BOUFFARD, P.; CALDERA, E.J.; CASH, E.; CAVANAUGH, A.; DENAS, O.; ELHAIK, E.; FAVÉ, M.J.; GADAU, J.; GIBSON, J.D.; GRAUR, D.; GRUBBS, K.J.; HAGEN, D.E.; HARKINS, T.T.; HELMKAMPF, M.; HU, H.; JOHNSON, B.R.; KIM, J; MARSH, S.E.; MOELLER, J.A.; MUÑOZ-TORRES, M.C.; MURPHY, M.C.; NAUGHTON, M.C.; NIGAM, S.; OVERSON, R.; RAJAKUMAR, R.; REESE, J.T.; SCOTT, J.J.; SMITH, C.R.; TAO, S.; TSUTSUI, N.D.; VILJAKAINEN, L.; WISSLER, L.; YANDELL, M.D.; ZIMMER, F.; TAYLO, J.; SLATER, S.C.; CLIFTON, S.W.; WARREN, W.C.; ELSIK, C.G.; SMITH, C.D.; WEINSTOCK, G.M.; GERARDO, N.M.; CURRIE, C.R. “The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle”. Plos Genet., **7**(2) :E1002007, 2011.

SCHEIN, C. H. "Production of soluble recombinant proteins in bacteria".
Biotechnology (New York; N. Y. 1983) **7**:1141–1149, 1989.

SCHECHTHER, I.; BERGER, A. On the size of the active site in peptidases. I Papin.
Biochemical and Biophysical Research Communications, **27**: 157-162, 1967.

STACK, C.M.; CAFFREY, C. R.; DONNELLY, S. M.; SESHADRI, A.; LOWTHER, J. F.; TORT, J. F.; COLLINS, P. R.; ROBINSON M. W.; XU, W.; MCKERROW J. H.; CRAIK, C. S.; GEIGER, S. R.; MARION, R.; BRINEN, L. S.; DALTON, J. P.
"Structural and functional relationships in the virulence-associated cathepsin L proteases of the parasitic liver fluke, *Fasciola hepatica*". J Biol Chem, **283**: 9896-9908, 2008.

SHENAI, B. R; SIJWALI, P. S; SINGH, A; ROSENTHAL, P. J.; "Enzyme Catalysis and Regulation: Characterizations of Native and Recombinant Falcipain-2, a Principal Trophozoite Cysteine Protease and Essential Cysteine Protease and Essential Hemoglobinase of *Plasmodium falciparum*". The Journal of Biological Chemistry, **275**(37): 29000-29010, 2000.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>> acesso em: Fev. 2018.

TAKAHASHI, N.; KURATA, S.; NATORI, S. "Molecular cloning of Cdna for the 29 kDa proteinase participating in decomposition of the larval fat body during metamorphosis of *Sarcophaga peregrina* (Flesh Fly)". Febs Lett., **334**: 153–157, 1993.

TANAKA, F. "Catalytic antibodies as designer proteases and esterases". Chem. Rev **102**:4885– 4906, 2002.

TERRA, W. R.; FERREIRA, C. "Biochemistry of Digestion". In Gilbert, L.I.; Iatov, K.; Gill, S. (Eds.). Comprehensive Molecular Insect Science, **4**:171-224, 2005.

TERRA, W.R; FERREIRA, C. "Insect digestive enzymes : properties, compartmentalization and function". Comparative Biochemistry and Physiology **109b**: 1–62, 1994

TERRA, W.R.; FERREIRA, C.; JORDAO, B.P.; DILLON, R.J. "Digestive enzymes". IN: Lehane, M.J., Billingsley, P.F. (Eds.), Biology of the Insect Midgut. Chapman and Hall, London, 1996. p. 153–194.

TOBBELL, D.A. ; MIDDLETON, B.J. ; RAINES, S. ; NEEDHAM, M.R. ; TAYLOR, I.W.; BEVERIDGE, J.Y. ; ABBOTT, W.M. "Identification of *in vitro* folding conditions for procathepsin S and cathepsin S using fractional factorial screens". Protein Expr. Purif., **24**:242–254, 2002.

TURK; B. "Targeting proteases: successes, failures and future prospects". Nat. Rev. Drug Discov., **5**: 785–799, 2006.

- TURK, V.; TURK, D. "Lysosomal cysteine proteases: facts and opportunities". *Embo J.*, **20**: 4629–4633, 2001.
- TURK, D.; GUNCAR, G. "Lysosomal cysteine proteases (cathepsins): promising drug targets". *Acta Crstallogr D Biol Crystallogr.*, D59: 203-312, 2003.
- TURK, B.; TURK, D.; TURK, V. "Protease signalling: the cutting edge". *The EMBO Journal*, **31**: 1630-1643, 2012a.
- TURK, V.; STOKA, V.; VASILJEVA, O.; RENKO, M.; SUN, T.; TURK, B.; TURK, D. "Cysteine cathepsins: from structure, function and regulation to new frontiers". *Biochimica et Biophysica Acta*, **1824**: 68–88, 2012b.
- TURK, D.; GUNCAR, G.; PODOBNIK, M.; TURK, B. "Revised definition of substrate binding sites of papain-like cysteine proteases". *Biol. Chem.* **379** :137–147, 1998.
- TURK, V.; BODE, W. "The cystatins: protein inhibitors of cysteine proteinases". *FEBS Lett.*, **285**: 213 – 219, 1991.
- UNIBRÁS AGROQUÍMICA. Técnico – Prejuízo causado pelas formigas cortadeiras. <http://www.unibras.com.br/tecnico-detalle/prejuizos-causados-pelas-formigas-cortadeiras/> acesso em Abril: 2018.
- VIEIRA, A. S.; BUENO, O. C.; CAMARGO-MATHIAS, M. I. "The functional morphology of the metapleural gland of the leaf-cutting ant *Atta laevigata* (Formicidae: *Attini*)". *Micron*, **41**:149–157, 2010.
- VILLAVERDE, A.; MAR CARRIO, M. "Protein aggregation in recombinant bacteria: biological role of inclusion bodies". *Biotechnol Lett*, **2**:1385-1395, 2003.
- WARD, P. S.; "ants". *Current Biology*, 16(5): R152, 2006.
- WENSELEERS, T. "The superorganism revisited". *Bio Science*, 59(8), 2009.
- WANG, S. X.; PANDEY, K. C.; SOMOZA, J. R.; SIJWALI, P. S.; KORTEMME, T.; BRINEN, L.S.; FLETTERICK, R. L.; ROSENTHAL, P. J.; MCKERROW, J.H. "Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease". *Proc Natl Acad Sci USA*, **103**: 11503-11508, 2006.
- WANG, K.M.; KUMAR, S.; CHENG, Y.S.; VENKATAGIRI, S.; YANG, A.H.; YEH, K.W. "Characterization of inhibitory mechanism and antifungal activity between group-1 and group-2 phytocystatins from taro (*Colocasia esculenta*)". *FEBS J.* **275**: 4980-4989, 2008.
- WATANABE, M.; YURA, A.; YAMANAKA, M.; KAMEI, K.; HARA, S; SUMIDA, M. "Purification and characterization of fibroinase, a cathepsin L-like cysteine proteinase, from the silk gland in the fourth instar *Bombyx mori* larva at the fourth molt period, stage D₂. *J. Insect Biotechnol. Sericol.*, 73:61-70, 2004.
- WEBER, NA. "Gardening ants, the *Attines*". Philadelphia: American Philosophical Society, 1972.

- WIEDERANDERS, B. "Structure-function relationships in class CA1 cysteine peptidase propeptides. *Acta Biochimica Polonica*, 50(3): 691-793, 2003.
- WONG, S.H. "Cloning of flavin reductase into Pet32a (+) expression vector lacking the thioredoxin a tag to study solubility of edta monooxygenase A in overexpression systems". *J. Exp. Microbiol. Immunol.*, **8**: 59-66, 2005.
- WOLFSON, J.L.; MURDOCK, L. L. "Potencial use os protease inhibitors for host plant resistance: a teste case". *Enviromental entomology*, 24(1): 52-77, 1995.
- WILSON, E.O. "The dominance of the social insects". IN: *Success and Dominance in Ecosystems*. Ecology Institute, Oldendorf_Luhe, Germany.1990. p. 1 -6.
- WILSON, E. O.; HOLLDOBLER, B. "The rise of the ants: A phylogenetic and ecological explanation". *Proc. Natl. Acad. Sci.*, 102(21): 7411-7414, 2005.
- ZANETTI, R.; CARVALHO, G.A.; SANTOS, A.; SOUZA SILVA, A.; GODOY, M.S. "*Manejo Integrado de Formigas Cortadeira*"; Universidade Federal de Lavras: Lavras, MG, Brazil, 2002; pp. 16
<<http://www.den.ufla.br/index.php/graduacao/get103-manejo-integrado-de-pragas-florestais>>. Acesso em: Março. 2018.
- ZANETTI, R.; ZANUNCIO, J.C.; MAYHÉ NUNES, A.J.; MEDEIROS, A.G.B.; SOUZA SILVA, A. " Combate sistemático de formigas cortadeiras com iscas granuladas, em eucaliptais com cultivo mínimo". *Revista Árvore*, **27**: 387–392, 2003.
- ZANETTI, R.; ZANUNCIO, J.C.; SANTOS, J.C.; SILVA, W.L.P.; RIBEIRO, G.T.; LEMES, P. G. An overview of integrated management of leaf-cutting ants (*Hymenoptera formicidae*) in brazilian forest plantations". *Forests* **5**: 439-454, 2014.
- YAMAMOTO, Y.; TAKAHASHI, S.Y. "Cysteine proteinase from Bombyx eggs: role in programmed degradation of yolk proteins during embryogenesis. *Comp. Biochem. Physiol.*, **B106**: 35–45, 1993.
- YAMAMOTO, Y.; WATABE, S.; KAGEYAMA, T.; TAKAHASHI, S.Y. "Pro region of *Bombyx Mori* cysteinase functions as an intramolecular chaperone to promote proper folding of the mature enzyme". *Arc insect Biochem Physiol.*, **42**: 167-178, 1999.
- YAMAGUCHI, H.; MIYAZAKI, M. "Refolding Techniques for Recovering Biologically Active Recombinant Proteins from Inclusion Bodies". *Biomolecules* , **4**: 235-251, 2014.
- YU, G.; SMITH, D.K.; ZHU, H.; GUAN, Y.; LAM, T.T.Y." ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data". *Methods in Ecology and Evolution*. 8(1): 28-36, 2017.

ANEXO A

Alguns espectros de MS/MS, resultados da análise da sequência de aminoácidos da proteína purificada AsCathL.

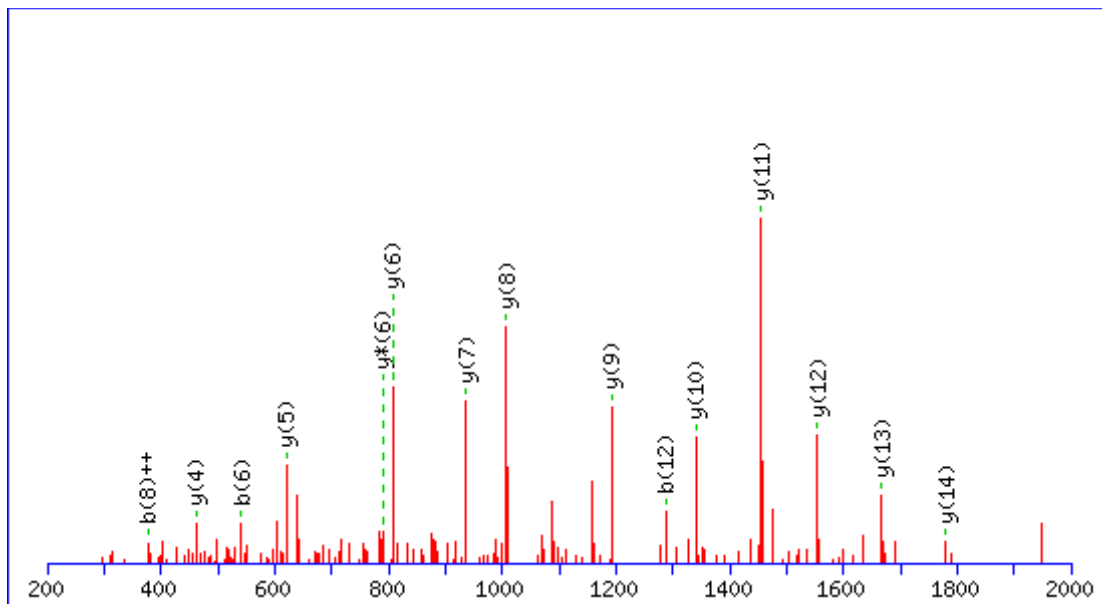
MS/MS Fragmentation of **ADGAILVDFWAEWCGPCK**

Monoisotopic mass of neutral peptide Mr(calc): 2093.9339

Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)

Match to Query 33055: 2093.954442 from(1047.984497,2+)

intensity(5166614.5000) rtinseconds(6745.9448) index(26686)



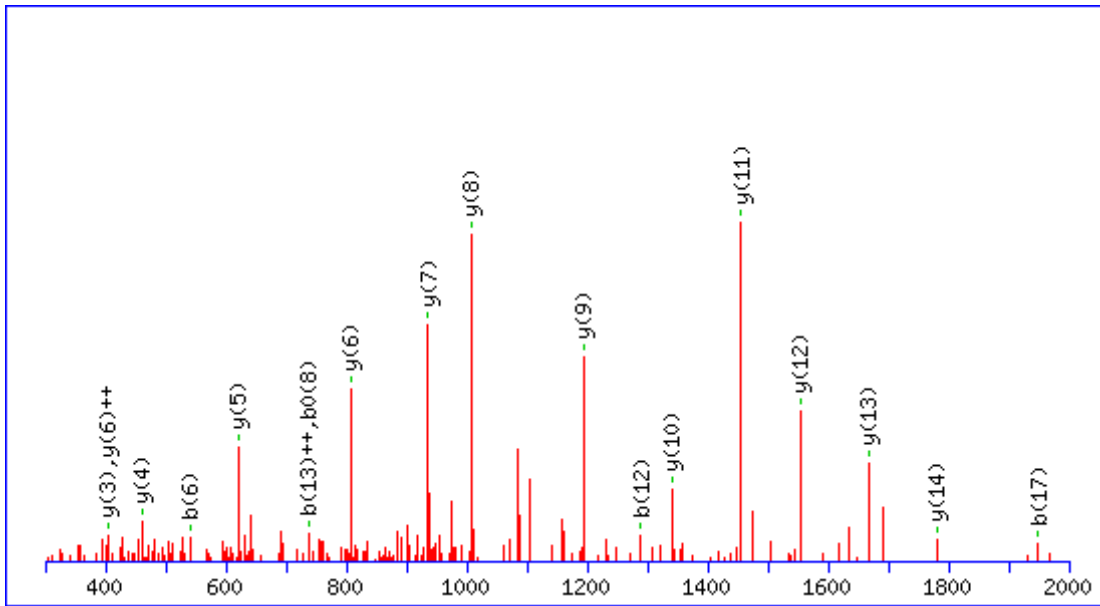
MS/MS Fragmentation of **ADGAILVDFWAEWCGPCK**

Monoisotopic mass of neutral peptide Mr(calc): 2093.9339

Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)

Match to Query 33077: 2093.955174 from(1047.984863,2+)

intensity(4576505.0000) rtinseconds(7086.2252) index(28360)



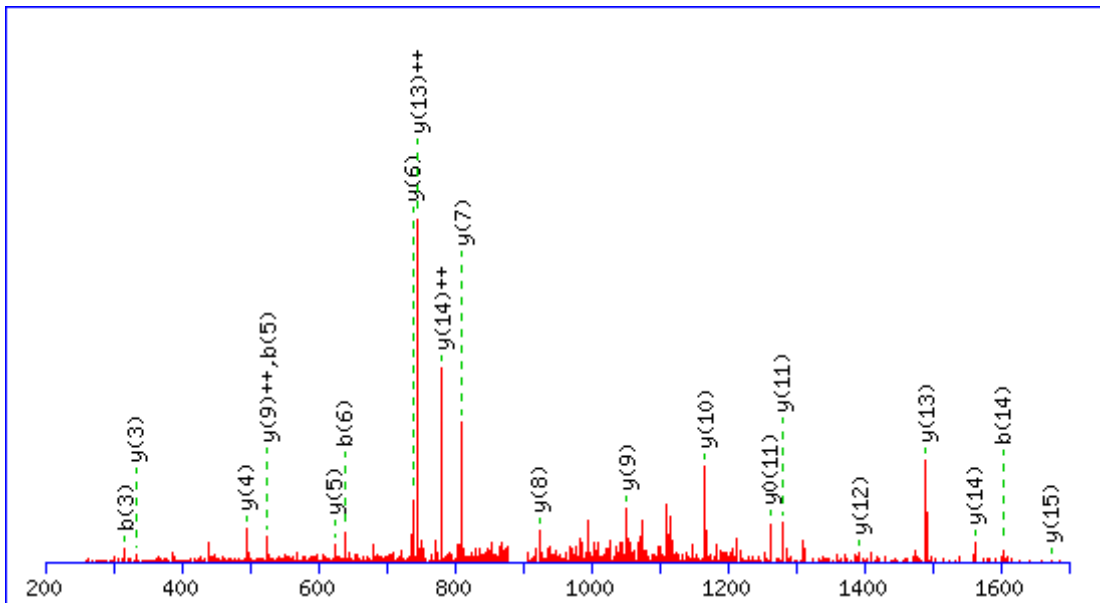
MS/MS Fragmentation of **MIAPILDEIADEYQGK**

Monoisotopic mass of neutral peptide Mr(calc): 1804.8917

Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)

Match to Query 24005: 1804.915990 from(903.465271,2+)

intensity(546538.5625) rtinseconds(5037.9767) index(17879)

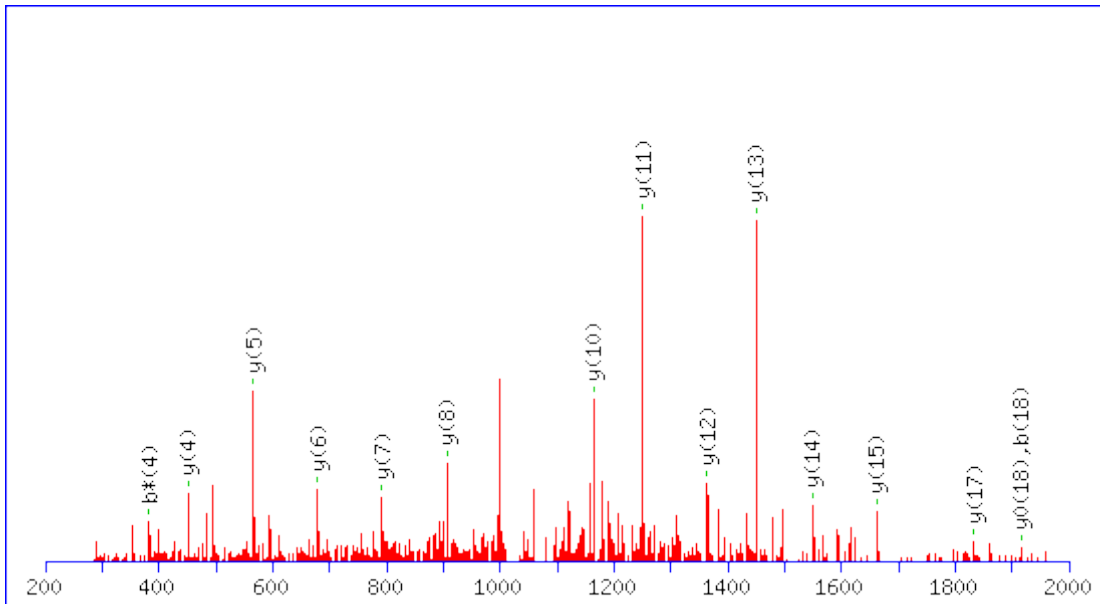


MS/MS Fragmentation of **QTGILVLSLEQNLI DCSGK**

Monoisotopic mass of neutral peptide Mr(calc): 2061.0412

Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)

Match to Query 32009: 2061.069432 from(1031.541992,2+)
intensity(475189.9375) rtinseconds(5877.9775) index(22212)



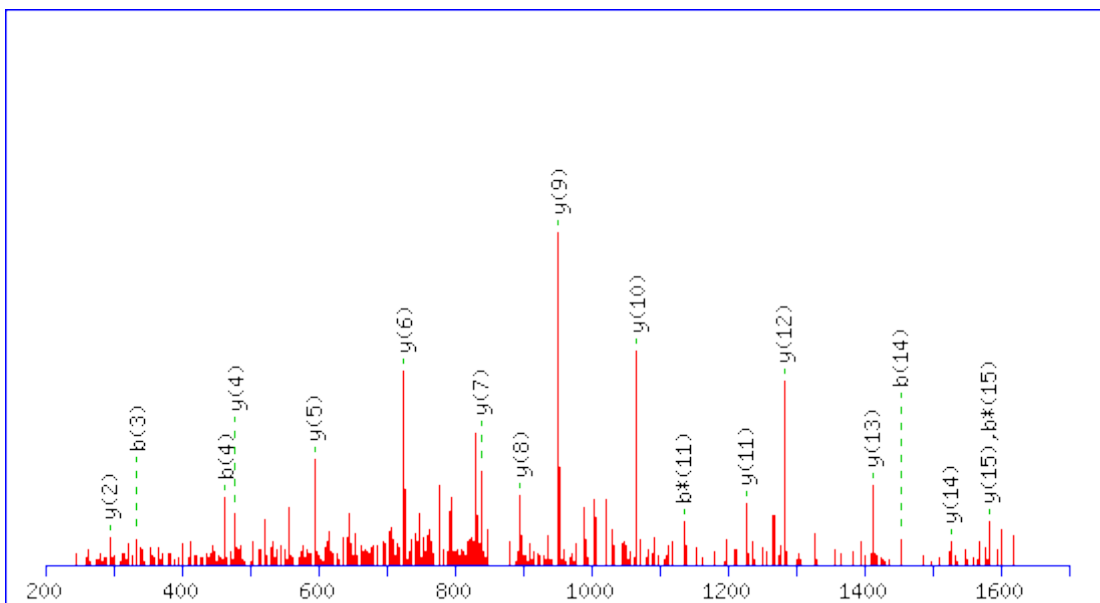
MS/MS Fragmentation of **YGNEGCNGGLMDNAFK**

Monoisotopic mass of neutral peptide Mr(calc): 1745.7137

Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)

Match to Query 21189: 1745.717992 from(873.866272,2+)

intensity(4622764.5000) rtinseconds(2530.5098) index(39421)



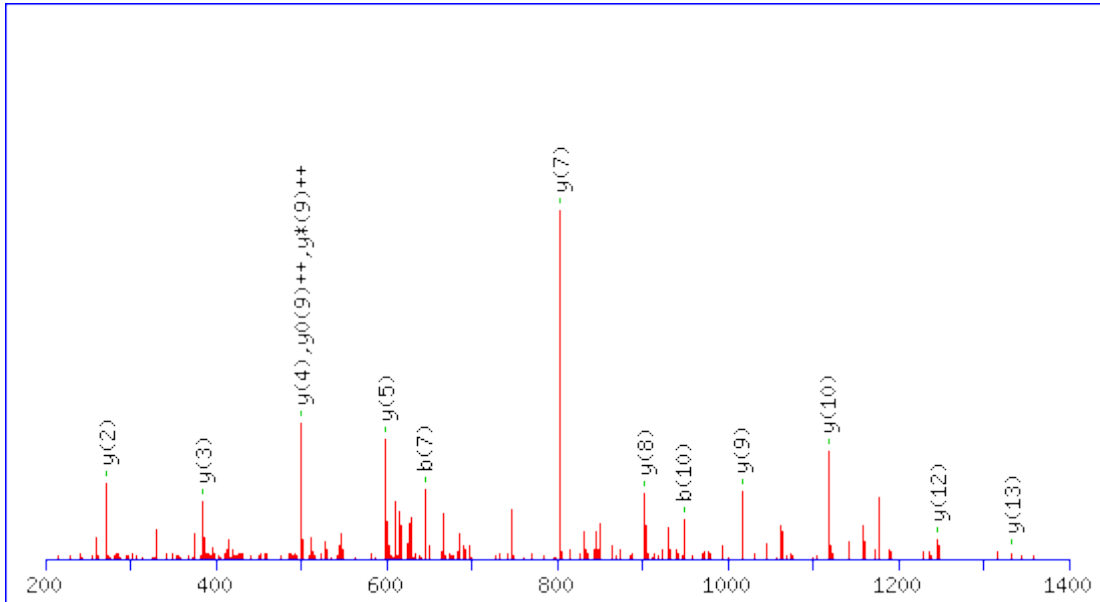
MS/MS Fragmentation of **NSGATDVGFDIPR**

Monoisotopic mass of neutral peptide Mr(calc): 1446.7103

Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)

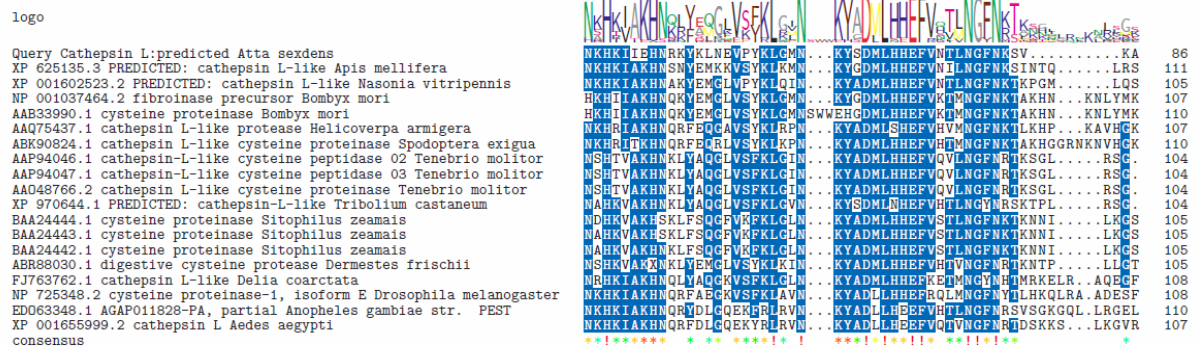
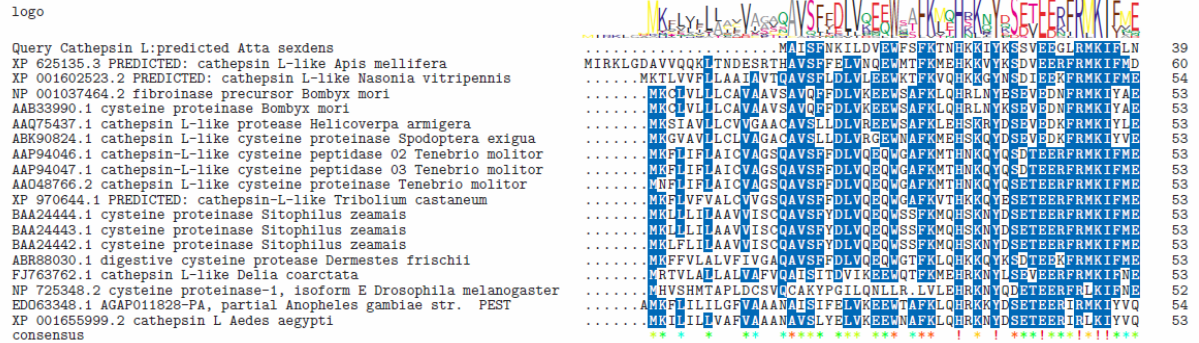
Match to Query 12157: 1446.714330 from(724.364441,2+)

intensity(149993648.0000) rtinseconds(2675.314) index(40159)



ANEXO B

Alinhamentos entre as sequências das formigas utilizando ClustalW.



logo

Query Cathepsin L:predicted Atta sexdens
 XP 625135.3 PREDICTED: cathepsin L-like Apis mellifera
 XP 001602523.2 PREDICTED: cathepsin L-like Nasonia vitripennis
 NP 001037464.2 fibroinase precursor Bombyx mori
 AAB33990.1 cysteine proteinase Bombyx mori
 AAQ75437.1 cathepsin L-like protease Helicoverpa armigera
 ABK90824.1 cathepsin L-like cysteine proteinase Spodoptera exigua
 AAP94046.1 cathepsin-L-like cysteine peptidase 02 Tenebrio molitor
 AAP94047.1 cathepsin-L-like cysteine peptidase 03 Tenebrio molitor
 AA048766.2 cathepsin L-like cysteine proteinase Tenebrio molitor
 XP 970644.1 PREDICTED: cathepsin-L-like Tribolium castaneum
 BAA24444.1 cysteine proteinase Sitophilus zeamais
 BAA24443.1 cysteine proteinase Sitophilus zeamais
 BAA24442.1 cysteine proteinase Sitophilus zeamais
 ABR88030.1 digestive cysteine protease Dermestes frischii
 FJ763762.1 cathepsin L-like Delia coarctata
 NP 725348.2 cysteine proteinase-1, isoform E Drosophila melanogaster
 ED063348.1 AGAP011828-PA, partial Anopheles gambiae str. PEST
 XP 001655999.2 cathepsin L Aedes aegypti
 consensus

logo

Query Cathepsin L:predicted Atta sexdens
 XP 625135.3 PREDICTED: cathepsin L-like Apis mellifera
 XP 001602523.2 PREDICTED: cathepsin L-like Nasonia vitripennis
 NP 001037464.2 fibroinase precursor Bombyx mori
 AAB33990.1 cysteine proteinase Bombyx mori
 AAQ75437.1 cathepsin L-like protease Helicoverpa armigera
 ABK90824.1 cathepsin L-like cysteine proteinase Spodoptera exigua
 AAP94046.1 cathepsin-L-like cysteine peptidase 02 Tenebrio molitor
 AAP94047.1 cathepsin-L-like cysteine peptidase 03 Tenebrio molitor
 AA048766.2 cathepsin L-like cysteine proteinase Tenebrio molitor
 XP 970644.1 PREDICTED: cathepsin-L-like Tribolium castaneum
 BAA24444.1 cysteine proteinase Sitophilus zeamais
 BAA24443.1 cysteine proteinase Sitophilus zeamais
 BAA24442.1 cysteine proteinase Sitophilus zeamais
 ABR88030.1 digestive cysteine protease Dermestes frischii
 FJ763762.1 cathepsin L-like Delia coarctata
 NP 725348.2 cysteine proteinase-1, isoform E Drosophila melanogaster
 ED063348.1 AGAP011828-PA, partial Anopheles gambiae str. PEST
 XP 001655999.2 cathepsin L Aedes aegypti
 consensus

logo

Query Cathepsin L:predicted Atta sexdens
 XP 625135.3 PREDICTED: cathepsin L-like Apis mellifera
 XP 001602523.2 PREDICTED: cathepsin L-like Nasonia vitripennis
 NP 001037464.2 fibroinase precursor Bombyx mori
 AAB33990.1 cysteine proteinase Bombyx mori
 AAQ75437.1 cathepsin L-like protease Helicoverpa armigera
 ABK90824.1 cathepsin L-like cysteine proteinase Spodoptera exigua
 AAP94046.1 cathepsin-L-like cysteine peptidase 02 Tenebrio molitor
 AAP94047.1 cathepsin-L-like cysteine peptidase 03 Tenebrio molitor
 AA048766.2 cathepsin L-like cysteine proteinase Tenebrio molitor
 XP 970644.1 PREDICTED: cathepsin-L-like Tribolium castaneum
 BAA24444.1 cysteine proteinase Sitophilus zeamais
 BAA24443.1 cysteine proteinase Sitophilus zeamais
 BAA24442.1 cysteine proteinase Sitophilus zeamais
 ABR88030.1 digestive cysteine protease Dermestes frischii
 FJ763762.1 cathepsin L-like Delia coarctata
 NP 725348.2 cysteine proteinase-1, isoform E Drosophila melanogaster
 ED063348.1 AGAP011828-PA, partial Anopheles gambiae str. PEST
 XP 001655999.2 cathepsin L Aedes aegypti
 consensus

logo

Query Cathepsin L:predicted Atta sexdens
 XP 625135.3 PREDICTED: cathepsin L-like Apis mellifera
 XP 001602523.2 PREDICTED: cathepsin L-like Nasonia vitripennis
 NP 001037464.2 fibroinase precursor Bombyx mori
 AAB33990.1 cysteine proteinase Bombyx mori
 AAQ75437.1 cathepsin L-like protease Helicoverpa armigera
 ABK90824.1 cathepsin L-like cysteine proteinase Spodoptera exigua
 AAP94046.1 cathepsin-L-like cysteine peptidase 02 Tenebrio molitor
 AAP94047.1 cathepsin-L-like cysteine peptidase 03 Tenebrio molitor
 AA048766.2 cathepsin L-like cysteine proteinase Tenebrio molitor
 XP 970644.1 PREDICTED: cathepsin-L-like Tribolium castaneum
 BAA24444.1 cysteine proteinase Sitophilus zeamais
 BAA24443.1 cysteine proteinase Sitophilus zeamais
 BAA24442.1 cysteine proteinase Sitophilus zeamais
 ABR88030.1 digestive cysteine protease Dermestes frischii
 FJ763762.1 cathepsin L-like Delia coarctata
 NP 725348.2 cysteine proteinase-1, isoform E Drosophila melanogaster
 ED063348.1 AGAP011828-PA, partial Anopheles gambiae str. PEST
 XP 001655999.2 cathepsin L Aedes aegypti
 consensus

logo

Query Cathepsin L:predicted Atta sexdens
 XP 625135.3 PREDICTED: cathepsin L-like Apis mellifera
 XP 001602523.2 PREDICTED: cathepsin L-like Nasonia vitripennis
 NP 001037464.2 fibroinase precursor Bombyx mori
 AAB33990.1 cysteine proteinase Bombyx mori
 AAQ75437.1 cathepsin L-like protease Helicoverpa armigera
 ABK90824.1 cathepsin L-like cysteine proteinase Spodoptera exigua
 AAP94046.1 cathepsin-L-like cysteine peptidase 02 Tenebrio molitor
 AAP94047.1 cathepsin-L-like cysteine peptidase 03 Tenebrio molitor
 AA048766.2 cathepsin L-like cysteine proteinase Tenebrio molitor
 XP 970644.1 PREDICTED: cathepsin-L-like Tribolium castaneum
 BAA24444.1 cysteine proteinase Sitophilus zeamais
 BAA24443.1 cysteine proteinase Sitophilus zeamais
 BAA24442.1 cysteine proteinase Sitophilus zeamais
 ABR88030.1 digestive cysteine protease Dermestes frischii
 FJ763762.1 cathepsin L-like Delia coarctata
 NP 725348.2 cysteine proteinase-1, isoform E Drosophila melanogaster
 ED063348.1 AGAP011828-PA, partial Anopheles gambiae str. PEST
 XP 001655999.2 cathepsin L Aedes aegypti
 consensus



logo

Query Cathepsin L:predicted Atta sexdens
 XP 625135.3 PREDICTED: cathepsin L-like Apis mellifera
 XP 001602523.2 PREDICTED: cathepsin L-like Nasonia vitripennis
 NP 001037464.2 fibroinase precursor Bombyx mori
 AAB33990.1 cysteine proteinase Bombyx mori
 AAQ75437.1 cathepsin L-like protease Helicoverpa armigera
 ABK90824.1 cathepsin L-like cysteine proteinase Spodoptera exigua
 AAP94046.1 cathepsin-L-like cysteine peptidase 02 Tenebrio molitor
 AAP94047.1 cathepsin-L-like cysteine peptidase 03 Tenebrio molitor
 AA048766.2 cathepsin L-like cysteine proteinase Tenebrio molitor
 XP 970644.1 PREDICTED: cathepsin-L-like Tribolium castaneum
 BAA24444.1 cysteine proteinase Sitophilus zeamais
 BAA24443.1 cysteine proteinase Sitophilus zeamais
 BAA24442.1 cysteine proteinase Sitophilus zeamais
 ABR88030.1 digestive cysteine protease Dermestes frischii
 FJ763762.1 cathepsin L-like Delia coarctata
 NP 725348.2 cysteine proteinase-1, isoform E Drosophila melanogaster
 ED063348.1 AGAP011828-PA, partial Anopheles gambiae str. PEST
 XP 001655999.2 cathepsin L Aedes aegypti
 consensus



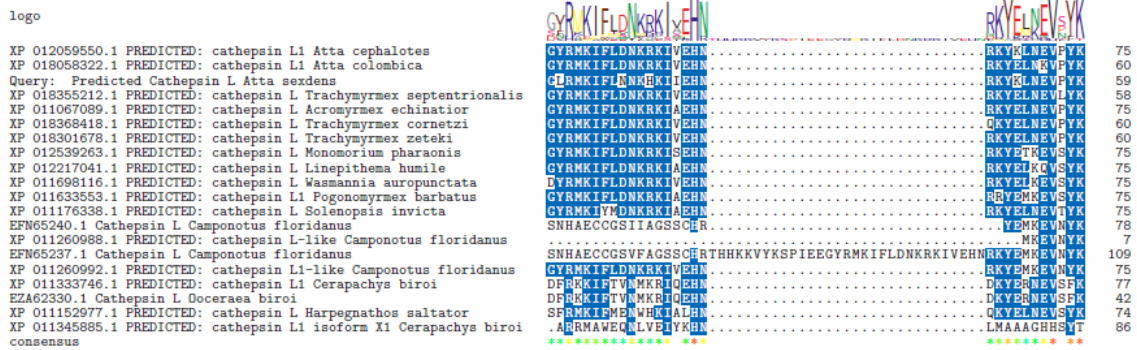
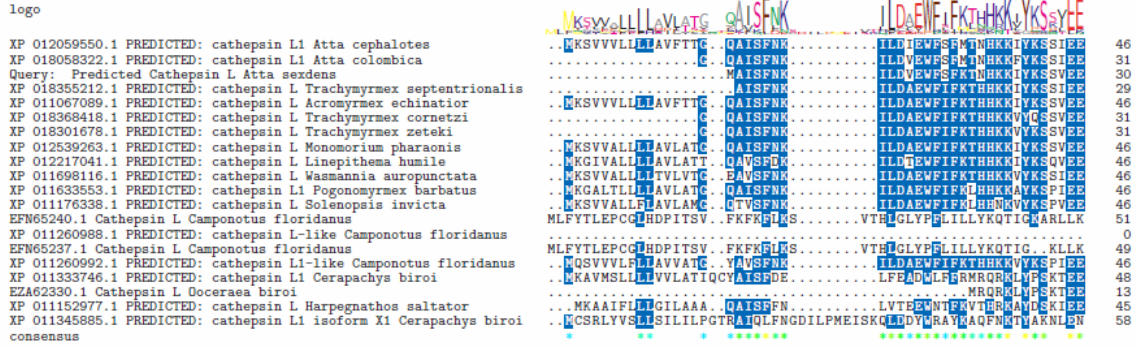
logo

Query Cathepsin L:predicted Atta sexdens
 XP 625135.3 PREDICTED: cathepsin L-like Apis mellifera
 XP 001602523.2 PREDICTED: cathepsin L-like Nasonia vitripennis
 NP 001037464.2 fibroinase precursor Bombyx mori
 AAB33990.1 cysteine proteinase Bombyx mori
 AAQ75437.1 cathepsin L-like protease Helicoverpa armigera
 ABK90824.1 cathepsin L-like cysteine proteinase Spodoptera exigua
 AAP94046.1 cathepsin-L-like cysteine peptidase 02 Tenebrio molitor
 AAP94047.1 cathepsin-L-like cysteine peptidase 03 Tenebrio molitor
 AA048766.2 cathepsin L-like cysteine proteinase Tenebrio molitor
 XP 970644.1 PREDICTED: cathepsin-L-like Tribolium castaneum
 BAA24444.1 cysteine proteinase Sitophilus zeamais
 BAA24443.1 cysteine proteinase Sitophilus zeamais
 BAA24442.1 cysteine proteinase Sitophilus zeamais
 ABR88030.1 digestive cysteine protease Dermestes frischii
 FJ763762.1 cathepsin L-like Delia coarctata
 NP 725348.2 cysteine proteinase-1, isoform E Drosophila melanogaster
 ED063348.1 AGAP011828-PA, partial Anopheles gambiae str. PEST
 XP 001655999.2 cathepsin L Aedes aegypti
 consensus



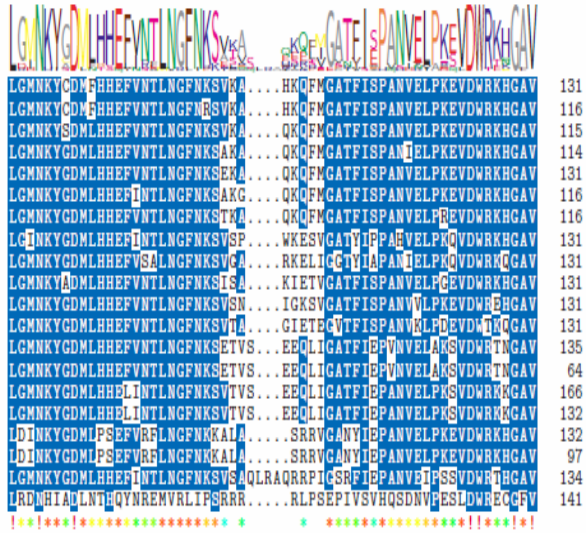
X non conserved
 X ≥ 50% conserved

Alinhamentos entre as sequências dos insetos utilizando o ClustalW



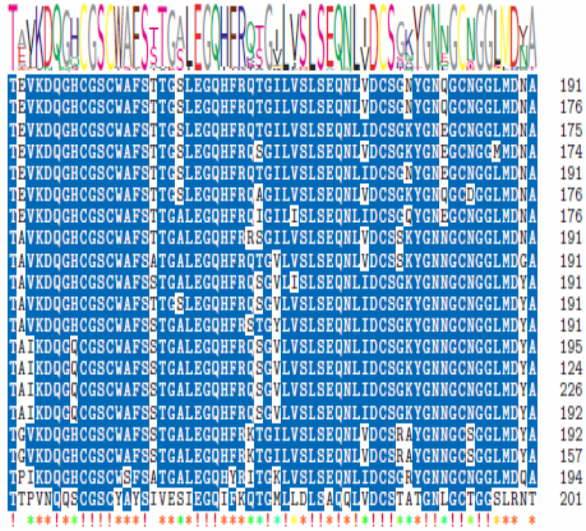
logo

XP 012059550.1 PREDICTED: cathepsin L1 Atta cephalotes
 XP 018058322.1 PREDICTED: cathepsin L1 Atta colombica
 Query: Predicted Cathepsin L Atta sexdens
 XP 018355212.1 PREDICTED: cathepsin L Trachymyrmex septentrionalis
 XP 011067089.1 PREDICTED: cathepsin L Acromyrmex echinator
 XP 018368418.1 PREDICTED: cathepsin L Trachymyrmex cornetzi
 XP 018301678.1 PREDICTED: cathepsin L Trachymyrmex zeteki
 XP 012539263.1 PREDICTED: cathepsin L Monomorium pharaonis
 XP 012217041.1 PREDICTED: cathepsin L Linepithema humile
 XP 011698116.1 PREDICTED: cathepsin L Wasmannia auropunctata
 XP 011633553.1 PREDICTED: cathepsin L Pogonomyrmex barbatus
 XP 011176338.1 PREDICTED: cathepsin L Solenopsis invicta
 EFN65240.1 Cathepsin L Camponotus floridanus
 XP 011260988.1 PREDICTED: cathepsin L-like Camponotus floridanus
 EFN65237.1 Cathepsin L Camponotus floridanus
 XP 011260992.1 PREDICTED: cathepsin L1-like Camponotus floridanus
 XP 011333746.1 PREDICTED: cathepsin L1 Cerapachys biroi
 EZA62330.1 Cathepsin L Ococeraea biroi
 XP 011152977.1 PREDICTED: cathepsin L Harpegnathos saltator
 XP 011345885.1 PREDICTED: cathepsin L1 isoform X1 Cerapachys biroi
 consensus



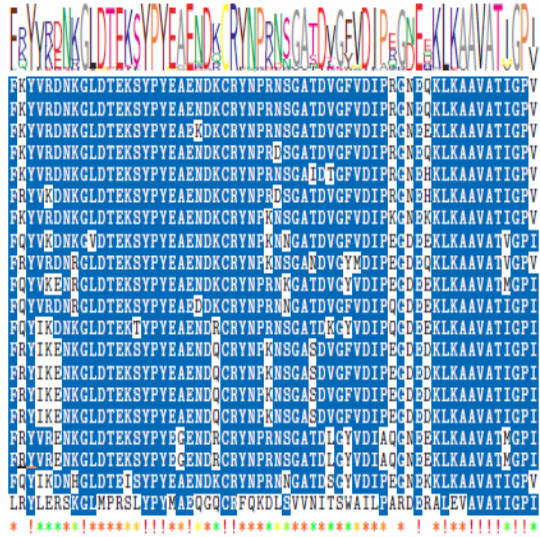
logo

XP 012059550.1 PREDICTED: cathepsin L1 Atta cephalotes
 XP 018058322.1 PREDICTED: cathepsin L1 Atta colombica
 Query: Predicted Cathepsin L Atta sexdens
 XP 018355212.1 PREDICTED: cathepsin L Trachymyrmex septentrionalis
 XP 011067089.1 PREDICTED: cathepsin L Acromyrmex echinator
 XP 018368418.1 PREDICTED: cathepsin L Trachymyrmex cornetzi
 XP 018301678.1 PREDICTED: cathepsin L Trachymyrmex zeteki
 XP 012539263.1 PREDICTED: cathepsin L Monomorium pharaonis
 XP 012217041.1 PREDICTED: cathepsin L Linepithema humile
 XP 011698116.1 PREDICTED: cathepsin L Wasmannia auropunctata
 XP 011633553.1 PREDICTED: cathepsin L1 Pogonomyrmex barbatus
 XP 011176338.1 PREDICTED: cathepsin L Solenopsis invicta
 EFN65240.1 Cathepsin L Camponotus floridanus
 XP 011260988.1 PREDICTED: cathepsin L-like Camponotus floridanus
 EFN65237.1 Cathepsin L Camponotus floridanus
 XP 011260992.1 PREDICTED: cathepsin L1-like Camponotus floridanus
 XP 011333746.1 PREDICTED: cathepsin L1 Cerapachys biroi
 EZA62330.1 Cathepsin L Ococeraea biroi
 XP 011152977.1 PREDICTED: cathepsin L Harpegnathos saltator
 XP 011345885.1 PREDICTED: cathepsin L1 isoform X1 Cerapachys biroi
 consensus



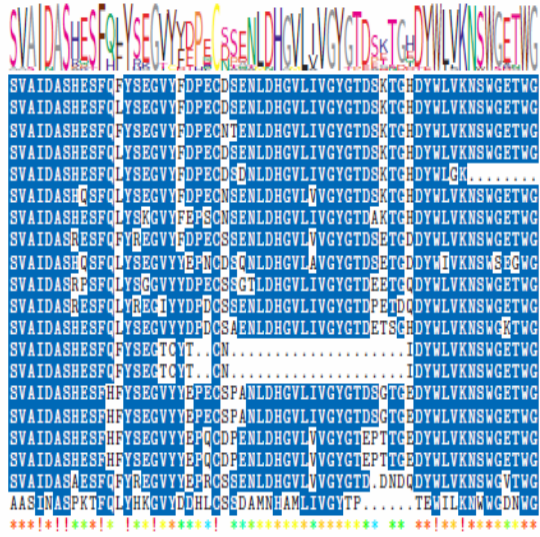
logo

XP 012059550.1 PREDICTED: cathepsin L1 Atta cephalotes
 XP 018058322.1 PREDICTED: cathepsin L1 Atta colombica
 Query: Predicted Cathepsin L Atta sexdens
 XP 018355212.1 PREDICTED: cathepsin L Trachymyrmex septentrionalis
 XP 011067089.1 PREDICTED: cathepsin L Acromyrmex echinator
 XP 018368418.1 PREDICTED: cathepsin L Trachymyrmex cornetzi
 XP 018301678.1 PREDICTED: cathepsin L Trachymyrmex zeteki
 XP 012539263.1 PREDICTED: cathepsin L Monomorium pharaonis
 XP 012217041.1 PREDICTED: cathepsin L Linepithema humile
 XP 011698116.1 PREDICTED: cathepsin L Wasmannia auropunctata
 XP 011633553.1 PREDICTED: cathepsin L1 Pogonomyrmex barbatus
 XP 011176338.1 PREDICTED: cathepsin L Solenopsis invicta
 EFM65240.1 Cathepsin L Camponotus floridanus
 XP 011260988.1 PREDICTED: cathepsin L-like Camponotus floridanus
 EFM65237.1 Cathepsin L Camponotus floridanus
 XP 011260992.1 PREDICTED: cathepsin L1-like Camponotus floridanus
 XP 011333746.1 PREDICTED: cathepsin L1 Cerapachys biroi
 EZA62330.1 Cathepsin L Ococeraea biroi
 XP 011152977.1 PREDICTED: cathepsin L Harpegnathos saltator
 XP 011345885.1 PREDICTED: cathepsin L1 isoform X1 Cerapachys biroi
 consensus



logo

XP 012059550.1 PREDICTED: cathepsin L1 Atta cephalotes
 XP 018058322.1 PREDICTED: cathepsin L1 Atta colombica
 Query: Predicted Cathepsin L Atta sexdens
 XP 018355212.1 PREDICTED: cathepsin L Trachymyrmex septentrionalis
 XP 011067089.1 PREDICTED: cathepsin L Acromyrmex echinator
 XP 018368418.1 PREDICTED: cathepsin L Trachymyrmex cornetzi
 XP 018301678.1 PREDICTED: cathepsin L Trachymyrmex zeteki
 XP 012539263.1 PREDICTED: cathepsin L Monomorium pharaonis
 XP 012217041.1 PREDICTED: cathepsin L Linepithema humile
 XP 011698116.1 PREDICTED: cathepsin L Wasmannia auropunctata
 XP 011633553.1 PREDICTED: cathepsin L1 Pogonomyrmex barbatus
 XP 011176338.1 PREDICTED: cathepsin L Solenopsis invicta
 EFM65240.1 Cathepsin L Camponotus floridanus
 XP 011260988.1 PREDICTED: cathepsin L-like Camponotus floridanus
 EFM65237.1 Cathepsin L Camponotus floridanus
 XP 011260992.1 PREDICTED: cathepsin L1-like Camponotus floridanus
 XP 011333746.1 PREDICTED: cathepsin L1 Cerapachys biroi
 EZA62330.1 Cathepsin L Ococeraea biroi
 XP 011152977.1 PREDICTED: cathepsin L Harpegnathos saltator
 XP 011345885.1 PREDICTED: cathepsin L1 isoform X1 Cerapachys biroi
 consensus



logo

XP 012059550.1 PREDICTED: cathepsin L1 Atta cephalotes
 XP 018058322.1 PREDICTED: cathepsin L1 Atta colombica
 Query: Predicted Cathepsin L Atta sexdens
 XP 018355212.1 PREDICTED: cathepsin L Trachymyrmex septentrionalis
 XP 011067089.1 PREDICTED: cathepsin L Acromyrmex echinator
 XP 018368418.1 PREDICTED: cathepsin L Trachymyrmex cornetzi
 XP 018301678.1 PREDICTED: cathepsin L Trachymyrmex zeteki
 XP 012539263.1 PREDICTED: cathepsin L Monomorium pharaonis
 XP 012217041.1 PREDICTED: cathepsin L Linepithema humile
 XP 011698116.1 PREDICTED: cathepsin L Wasmannia auropunctata
 XP 011633553.1 PREDICTED: cathepsin L1 Pogonomyrmex barbatus
 XP 011176338.1 PREDICTED: cathepsin L Solenopsis invicta
 EFM65240.1 Cathepsin L Camponotus floridanus
 XP 011260988.1 PREDICTED: cathepsin L-like Camponotus floridanus
 EFM65237.1 Cathepsin L Camponotus floridanus
 XP 011260992.1 PREDICTED: cathepsin L1-like Camponotus floridanus
 XP 011333746.1 PREDICTED: cathepsin L1 Cerapachys biroi
 EZA62330.1 Cathepsin L Ococeraea biroi
 XP 011152977.1 PREDICTED: cathepsin L Harpegnathos saltator
 XP 011345885.1 PREDICTED: cathepsin L1 isoform X1 Cerapachys biroi
 consensus

