

Natássia Caroline Resende Corrêa

**Clonagem e expressão de
protease mitogênica de
*Carica candamarcensis***

Belo Horizonte - MG
2009

Natássia Caroline Resende Corrêa

Clonagem e expressão de protease mitogênica de *Carica candamarcensis*

Dissertação de Mestrado apresentada ao Departamento de Bioquímica e Imunologia do Instituto de Ciências Biológicas, da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Bioquímica e Imunologia.

Orientador: Carlos Edmundo Salas Bravo

Belo Horizonte - MG
2009

Aos meus pais e irmãos.

Agradecimentos

Agradeço a Deus, pelas oportunidades que Ele me deu durante o mestrado, pelo consolo nos momentos difíceis e por cada vitória que alcancei.

Aos meus pais, por todo apoio e amor que me deram. Mesmo estando distantes, foram muito importantes para essa caminhada. Muito obrigada também pelo incentivo de sempre buscar pelo o que é melhor para mim.

Às minhas irmãs por sempre me incentivarem em tudo o que faço, pelo apoio, pelo amor, pelas risadas e pela compreensão e conselhos nas horas mais difíceis desses dois anos.

À minha família, em especial às minhas tias Fátima e Lúcia, pela acolhida em suas casas e pelo apoio incondicional.

Ao meu orientador Carlos Salas, pelos incontáveis ensinamentos durante esses dois anos. Com certeza, sem ele não teria conseguido o que tenho hoje. Muito obrigada por me ajudar nas questões profissionais e pessoais, sempre me lembrarei do que tem feito por mim.

Ao Danilo, meu namorado, pela imensa compreensão e carinho nestes dois anos em que estamos em Belo Horizonte. Muito obrigada pela força em tudo que faço e pelo incentivo para sempre ser melhor. Sem você aqui eu não teria conseguido!

Às queridas amigas Raquel, Isabela e Gabi. Sem vocês o mestrado simplesmente não teria graça. Muito obrigada pela ajuda nos experimentos, mas principalmente pelas várias horas de descontração, especialmente nos momentos difíceis que passei durante o mestrado. Vocês são muito especiais, sei que sempre poderei contar com vocês.

Aos amigos do laboratório de Biologia Molecular de Produtos Naturais, vocês significam muito na minha vida. Muito obrigada pelos momentos engraçadíssimos, pelo apoio em todos os experimentos, pelo silêncio quando era necessário... Marco

Túlio, Raphael, Henrique, Cacai e Andréia ... todos vocês me ensinaram inúmeras coisas e sou muito grata a vocês.

Ao Tércio e ao Carlos Gustavo, pessoas essenciais para a realização desse trabalho. Vocês me ajudaram a superar todas as barreiras que apareceram no caminho, serei eternamente grata.

Aos amigos de outros laboratórios Bruna, Débora, Aline, André, Cíntia, Kádima e Lílian pela ajuda técnica e convivência mais que agradável.

Aos colegas Caroline, Jane, Daniel, Rodrigo, William, Higor e Joacir, pela ajuda nos experimentos.

À Neusa!, minha inseparável companheira.

Ao Jamil, pelo imenso altruísmo, pelos incontáveis reagentes emprestados, pelas dicas infalíveis e por sempre me compreender nas horas difíceis.

Aos amigos do LSAT, especialmente ao Mito (Ana Cândida), Celso, Cinthia, Sandrinha, Fernanda, Kátia e à Luciana, pessoa sem a qual nada acontece. Muito obrigada Lulu.

A Louisa e ao Rubens, pela amizade, por me receberem em sua casa com tanto carinho e me apoiarem nos momentos difíceis.

Aos amigos de Bases, em especial à Paula e à Ju pela amizade.

Aos meus amigos Samuel, Rub, Taís, Amanda, Elias, Flavi, Flávia, Gabriele, Dodo e Di, que mesmo distante, se fizeram extremamente presentes em minha vida.

Aos professores do ICB, especialmente à Prof. Míriam Tereza, Prof. Alfredo Góes, Prof. Evanguedes e ao Prof. Ronaldo Nagem por toda a ajuda que me ofereceram nesses dois anos.

Ao CNPq por financiar esse trabalho.

Resumo

As cisteíno-proteases são encontradas em vários organismos e se destacam por apresentar funções essenciais em parasitos, mamíferos e plantas. Várias plantas que são fontes dessas cisteíno-proteases pertencem à família Caricaceae, que tem como membro a espécie *Carica candamarcensis*. Uma das proteínas isoladas desse organismo, denominada CMS2MS2, é composta por 214 aminoácidos, possui atividade proteolítica e comprovada ação mitogênica em células de mamíferos. O objetivo desse trabalho foi produzir CMS2MS2 recombinante para posteriores estudos da atividade biológica e caracterização estrutural. Para a clonagem desse gene, foi utilizado um par de “primers” degenerados baseados na seqüência aminoacídica N-terminal e C-terminal da proteína CMS2MS2, cuja estrutura foi recentemente elucidada. Amplicons de aproximadamente 650 e 500 bp foram obtidos do cDNA, gerado por transcrição reversa a partir do RNA da folha da planta. A clonagem inicial do fragmento de 650 bp foi realizada no vetor pCR 2.1, que por sua vez foi transformado em bactérias *Escherichia coli* TOP10. Os clones positivos foram selecionados e a presença do fragmento de interesse verificada através de digestão, PCR e seqüenciamento do inserto. Foi possível observar pela análise dos clones F2 e E3 a existência de possíveis isoformas de cisteíno-proteases em *C. candamarcensis*. O inserto do clone F2 foi escolhido e em seguida excisado do plasmídeo pCR 2.1 com as enzimas XhoI e HindIII para sub-clonagem no vetor pET28a (+) TEV entre os sítios Sall e HindIII. O perfil de expressão dos clones em bactérias BL21-DE3 induzidas com 0,5 mM IPTG demonstrou uma proteína de aproximadamente 30 kDa, como esperado após a sub-clonagem, presente somente na fração insolúvel do lisado bacteriano. Testes de atividade mitogênica da proteína recombinante mostrou que essa proteína possui efeito proliferativo sobre fibroblastos L929 nas concentrações de 0,4 e 1 nM, enquanto que em osteobalstos obtidos de cultura primária induz mitogênese na concentração de 1 nM, faixas de concentração nas quais a enzima não-recombinante exibe atividade farmacológica.

Palavras-chave: cisteíno-protease, CMS2MS2, clonagem molecular, expressão, efeito mitogênico.

Abstract

Cysteine proteinases are found in many organisms and play essential functions in parasites, mammals and plants. *Carica candamarcensis*, a member of the Caricaceae family, contains highly active cysteine proteinases in its latex. One of the proteins isolated from this specie was named CMS2MS2, it contains 214 aminoacids, exhibits proteolytic activity and shows mitogenic effect on mammal cells. The objective of this study was to produce recombinant CMS2MS2 to further study the biological activity and structure characterize this enzyme. To clone this gene a degenerate primer pair was derived from the recently elucidated N- and C-terminal sequence of CMS2MS2,. Amplicons of about 650 bp and 500 bp were obtained by PCR from the cDNA, generated by reverse transcription using total RNA from leaf. Initial cloning of the 650 bp fragment was done into plasmid pCR 2.1, which was transformed into *Escherichia coli* TOP10. Positive clones were confirmed by endonuclease digestion, specific PCR and sequencing. It was shown that F2 and E3 clones could be isoforms of cysteine-proteases in *C. candamarcensis*. The F2 clone was chosen and excised from pCR 2.1 with XhoI and HindIII enzymes for sub-cloning into pET28a(+)-TEV vector, between Sall e HindIII sites. The expression profile in BL21-DE3 bacteria induced with 0.5 mM IPTG showed a protein with approximately 30 kDa, as expected. The protein is only present in the insoluble fraction of the bacterial lysate. The mitogenic activity of the recombinant protein showed that the protein has proliferative effect on L929 fibroblasts at 0.4 e 1 nM concentrations, while in primary cultured osteoblasts it induces mitogenesis at 1 nM concentration which is within the pharmacological range exhibited by the non-recombinant enzyme.

Key-words: cysteine-protease, CMS2MS2, molecular cloning, expression, mitogenic effect.

Lista de figuras

Figura 1: Mapa do vetor pCR 2.1, mostrando os sítios de restrição (TA Cloning kit, Invitrogen).	31
Figura 2: Diagrama de Venn, utilizado pelo programa Clustalw EBI para verificar a natureza das substituições de aminoácidos.	38
Figura 3: Mapa do vetor pET28a(+)-TEV, mostrando os sítios de restrição.....	39
Figura 4: Eletroforese em gel de agarose/formaldeído de RNA total extraído de folhas de <i>C. candamarcensis</i>	46
Figura 5: Eletroforese de fragmento amplificado a partir de cDNA utilizando os primers CMS2MS2	47
Figura 6: Eletroforese de DNA plasmidial pCR 2.1 dos possíveis clones de CMS2MS2, F2 e E3. Gel de agarose 1% corado com brometo de etídeo	48
Figura 7: Sequência de nucleotídeos do clone F2, obtida em sequenciador automático 3.100 Genetic Analyzer (Applied Biosystems), utilizando os primers M13 “forward” ou “reverse”	49
Figura 8: Sequência de nucleotídeos do clone E3, obtida em sequenciador automático 3.100 Genetic Analyzer (Applied Biosystems), utilizando os primers M13 “forward” ou M13 “reverse”	49
Figura 9: Sequência da proteína hipotética da região do inserto do clone F2 correspondente à região codificante de CMS2MS2	50
Figura 10: Sequência da proteína hipotética da região do inserto do clone E3 correspondente à região codificante de CMS2MS2	50
Figura 11: Alinhamento das sequências de aminoácidos traduzidas <i>in silico</i> dos clones F2 e E3	53
Figura 12: Alinhamento das sequências de aminoácidos obtidas dos clones F2 e E3 com a protease CMS2MS2..	54
Figura 13: Alinhamento das sequências de aminoácidos de F2 (traduzida <i>in silico</i> a partir da sequência de DNA) e CMS2MS3	55
Figura 14: Esquema representativo do resultado da sub-clonagem em pET28a(+)-TEV	57

Figura 15: Eletroforese em gel de agarose 1% de fragmento amplificado a partir do inserto F2, sub-clonado em vetor pET28a(+)-TEV, utilizando os primers específicos para CMS2MS2. Gel de agarose 1% corado com brometo de etídeo.....	58
Figura 16: Eletroforese em gel SDS-PAGE da fração solúvel da expressão da proteína recombinante presente no clone 1	59
Figura 17: Eletroforese em gel SDS-PAGE da fração insolúvel da expressão da proteína recombinante presente no clone.	59
Figura 18: Eletroforese em gel SDS-PAGE da fração insolúvel da expressão da proteína recombinante presente nos clones 1 e 3.....	60
Figura 19: Sequência de aminoácidos traduzida <i>in silico</i> do clone F2, mostrando as duas caudas de histidina, o “stop” códon, e em preto o início e final da sequência codificante de CMS2MS2.....	61
Figura 20: Eletroforese em SDS-PAGE mostrando as frações eluídas da coluna de afinidade.....	63
Figura 21: Índice de proliferação de células L929 após tratamento com CMS2MS2 não-recombinante e proteína recombinante.....	64
Figura 22: Índice de proliferação de osteoblastos após tratamento com CMS2MS2 natural e CMS2MS2 recombinante	65

Lista de tabelas

Tabela 1: Descrição dos cinco maiores “scores” de identidade da sequência de nucleotídeos F2 em relação à sequências depositadas em banco de dados, utilizando o algoritmo “blastn” do servidor NCBI.	51
Tabela 2: Descrição dos cinco maiores “scores” de identidade da sequência de nucleotídeos E3 em relação à sequências depositadas em banco de dados, utilizando o algoritmo “blastn” do servidor NCBI.	51
Tabela 3: Descrição dos cinco maiores “scores” de identidade da sequência de aminoácidos de F2 com sequências depositadas em banco de dados, utilizando o programa “blastx” do NCBI.....	52
Tabela 4: Descrição dos cinco maiores “scores” de identidade da sequência de aminoácidos de E3 com sequências depositadas em banco de dados, utilizando o programa “blastx” do NCBI.....	52
Tabela 5: Substituições de aminoácidos da proteína codificada pelo clone F2 em relação a CMS2MS2, mostrando o tipo de substituição e a natureza da cadeia lateral quando a substituição foi conservativa ou semi-conservativa.	56
Tabela 6: Parâmetros bioquímicos de CMS2MS2 e da proteína recombinante F2 expressa.....	62

Lista de abreviaturas

%	porcentagem
°C	graus Célsius ou Centígrados
bp	pares de bases
cDNA	ácido desóxi-ribonucleico complementar
Ci	Curie
cm	centímetro (10^{-2} m)
CPM	cintilações por minuto
C-terminal	região carbóxi-terminal da proteína
D	na sequência de primers representa A, G ou T
Da.	Dalton
DNA	“desoxyribonucleic acid” (ácido desoxi-ribonucléico)
dNTP	desóxi-nucleotídeos
<i>E. coli</i>	<i>Escherichia coli</i>
et al.	et all (e demais colaboradores)
H	na sequência de primers representa A, C ou T
h	hora
kb	Kilobases
kDa.	KiloDalton (10^3 Da.)
L	litro
M	concentração molar (moles/L)
m	metro
mA	miliAmpère
mCi	miliCurie (10^{-3} Ci)
μCi	microCurie (10^{-6} Ci)
mg	miligrama (10^{-3} g)
μg	micrograma (10^{-6} g)
min	minuto
ml	mililitro (10^{-3} L)
μl	microlitro (10^{-6} L)
mM	concentração milimolar (10^{-3} moles/L)
μM	concentração micromolar (10^{-6} moles/L)

μm	micrômetro (10^{-6} m)
mRNA	ácido ribonucléico mensageiro
MW	“molecular weight” (peso molecular)
N	na sequência de primers representa A, G, C ou T
ng	nanograma (10^{-9} g)
nM	concentração nanomolar (10^{-9} moles/L)
nm	nanômetro (10^{-9} m)
N-terminal	região amino-terminal da proteína
OD	“optical density” (densidade óptica)
P.A.	“Pro Analysis” (padrão ACS- “American Chemical Society”)
PCR	“polymerase chain reaction” (reação em cadeia da polimerase)
pH	potencial hidrogeniônico
pI	potencial isoelétrico
q.s.p.	quantidade suficiente para
R	na sequência de primers representa A ou G
RNA	“ribonucleic acid” (ácido ribonucléico)
rpm	rotações por minuto
RT	“reverse transcriptase” (transcriptase reversa)
s	segundo
U	unidade
V	Volts
W	Watts
Y	na sequência de primers representa C ou T

As demais abreviaturas e siglas serão explicadas quando forem citadas pela primeira vez no texto.

Sumário

1. Introdução	1
1.1 - A família Caricaceae.....	1
1.2 - Prováveis funções do látex.....	2
1.3 - As cisteíno-proteases	3
1.3.1 – Classificação e distribuição na natureza	3
1.3.2 – Características estruturais de cisteíno-proteases da família C1, clã CA..	4
1.3.3 – Cisteíno-proteases de <i>Carica candamarcensis</i> e suas propriedades biológicas	6
1.4 – Amplificação, clonagem e expressão de genes de cisteíno-proteases de Caricaceae.....	10
1.5 - O ciclo celular em eucariotos.....	11
1.6 - Fatores de crescimento celular.....	13
1.7 - Proteases como fatores de proliferação	15
3- Materiais e Métodos.....	19
3.1- Materiais.....	19
3.1.1- Reagentes Químicos.....	19
3.1.2- Soluções	20
3.1.3- Biológicos.....	27
3.2 - Métodos.....	27
3.2.1 - Extração de RNA	27
3.2.2- Síntese de cDNA.....	29
3.3.3 - Amplificação do gene de CMS2MS2 a partir de cDNA	29
3.3.4 - Análise e quantificação do produto de PCR.....	30
3.3.5 - Ligação em pCR 2.1	31
3.3.6 - Preparo de bactérias quimiocompetentes.....	32
3.3.7 - Transformação de bactérias quimiocompetentes	32
3.3.8 - Seleção de clones positivos.....	33
3.3.9 - Extração de DNA plasmidial (“mini-prep”).....	33
3.3.10 - Verificação da presença do inserto.....	34
3.3.11 - Sequenciamento dos clones positivos	35
3.3.12 - Análises <i>in silico</i>	37

3.3.13 - Purificação do inserto e obtenção do vetor pET28a(+)-TEV	38
3.3.14 - Ligação do inserto ao vetor pEt28a(+)-TEV e transformação em bactérias <i>E. coli</i> BL21-DE3	40
3.3.15 - Expressão de CMS2MS2 recombinante	41
3.3.16 - Eletroforese em Gel de Poliacrilamida Desnaturante:.....	42
3.3.17 - Análise de parâmetros bioquímicos	43
3.3.18 - Purificação por cromatografia de afinidade ao níquel	43
3.3.19 - Ensaio de atividade mitogênica com CMS2MS2 recombinante	44
4. Resultados	46
4.1 - Amplificação de CMS2MS2 e clonagem em vetor pCR 2.1.....	46
4.2 - PCR, sequenciamento dos clones e análise <i>in silico</i>	48
4.3 - Sub-clonagem no vetor pET28a(+)-TEV	56
4.4 - Expressão da proteína recombinante	58
4.5 - Análise dos parâmetros bioquímicos.....	61
4.6 - Purificação da proteína recombinante	62
4.7 - Ensaio de atividade biológica com CMS2MS2 recombinante.....	63
5. Discussão.....	67
6. Conclusões.....	77
7. Referências Bibliográficas.....	79

1. Introdução

1.1 - A família Caricaceae

A Caricaceae é uma família de dicotiledôneas que possui seis gêneros, um deles presente na África Equatorial e os outros cinco distribuídos do centro ao sul da América (Kyndt *et al.*, 2005). Esse grupo contém diversas espécies frutíferas produtoras de látex, material biológico de alto interesse comercial por ser fonte de enzimas proteolíticas com diversas aplicações industriais e farmacológicas (Salas *et al.*, 2008)

Dentre as espécies predominantes na América Central e Sul, a principal representante é a planta *Carica papaya*, da qual uma mistura de enzimas, denominada papaína, é extraída e utilizada em vários processos industriais como amaciamento de carnes, curtimento de couros, clarificação de cervejas e solubilização de proteínas insolúveis (Duke & duCellier, 1993). Uma das enzimas isoladas dessa mistura, também chamada papaína, é a proteína mais bem estudada do grupo de cisteíno-proteases (Silva *et al.*, 1997).

Carica candamarcensis, outra espécie da família Caricaceae, é nativa da América do Sul e é encontrada principalmente na região Andina, da Colômbia até o Chile. Ela é encontrada predominantemente em altitudes superiores a 1000 metros, com temperatura média variando entre 12 e 18 °C, em clima sub-úmido. É conhecida popularmente como “mountain papaya” (“mamão da montanha”) ou “highland papaya” (Walraevens *et al.*, 1999). Pode chegar até 10 metros de altura, possuindo uma coroa de folhas na região apical do tronco. O fruto é elipsóide, amarelo quando maduro, mede de 5 a 15 cm de comprimento e somente pode ser consumido após cozimento, devido a seu sabor azedo e adstringente (Ortega & Pino, 1997).

Recentemente, uma nova classificação dos gêneros foi sugerida, em que *Carica papaya* seria a única espécie do gênero *Carica*, enquanto as outras espécies americanas se enquadrariam no gênero *Vasconcellea*, entre elas *Carica candamarcensis* que também é chamada de *Vasconcellea cundinamarcensis* (Badillo, 2000). Além disso, vários estudos filogenéticos analisaram espécies da família Caricaceae (Van Droogenbroeck *et al.*, 2004; Kyndt *et al.*, 2007), buscando

entender principalmente as mudanças genéticas causadas pela hibridação entre as espécies comuns da América do Sul. Estas pesquisas demonstraram que houve troca de material genético entre várias espécies de *Vasconcellea* e que é possível que *C. candamarcensis* e *V. stipulata* tenham dado origem à espécie *V. x heilbornii*, ainda que essa hipótese não tenha sido confirmada (Van Droogenbroeck *et al.*, 2004; Kyndt *et al.*, 2005). Essas hibridações indicam que alguns genes podem estar sendo herdados frequentemente de outras espécies e que sub-especializações, duplicações gênicas e o próprio tempo de evolução produziram nessas espécies proteínas ligeiramente diferentes em sequência e função (Van Droogenbroeck *et al.*, 2004; Kyndt *et al.*, 2005).

1.2 - Prováveis funções do látex

O látex de Caricaceae é uma substância pegajosa, de cor branca a transparente e é produzido em quase toda a planta, principalmente em partes aéreas (Moussaoui *et al.*, 2001; Kyndt *et al.*, 2007). As espécies de Caricaceae produzem e armazenam o látex em células chamadas laticíferos, que são interconectados por estruturas ramificadas, os canais laticíferos, formando uma complexa rede de tubos pela planta (Moussaoui *et al.*, 2001). Quando a planta sofre uma injúria, o látex se acumula no local exposto e extravasa por aquela abertura. Em alguns minutos um coágulo se forma, selando a área ferida, evitando a entrada de patógenos e a perda de fluidos (Silva *et al.*, 1997; Azarkan *et al.*, 2004).

No látex também são encontradas substâncias como as quitinases, inibidores de serino-proteases e cisteíno-proteases, que participam na defesa da planta contra patógenos e predadores.

As quitinases podem estar ligadas à ação anti-fúngica e também na degradação da quitina presente no aparelho bucal de muitos insetos. Os inibidores de serino-proteases estariam atuando nas enzimas do aparelho digestório de insetos e outros organismos, tornando as plantas tóxicas ou não-digestivas (Azarkan *et al.*, 2004; Konno *et al.*, 2004).

Já as cisteíno-proteases, abundantes no látex de Caricaceae, parecem estar envolvidas no processo de defesa contra larvas de insetos, uma vez que Konno e colaboradores (2004) demonstraram que látex de *C. papaya* inibia o crescimento e matava esses organismos, num processo dependente da ação proteolítica de

cisteíno-proteases. Além disso, foi demonstrado que ocorrem clivagens sucessivas e ordenadas de peptídeos durante o extravasamento de látex em *C. papaya* (Silva *et al.*, 1997), indicando que esse evento possa estar relacionado com o fechamento da abertura da ferida (Silva *et al.*, 1997). A liberação, ativação e processamento de peptídeos em sequência e formação de coágulo no local da ferida sugerem um mecanismo de defesa similar ao encontrado em tecidos animais (Silva *et al.*, 1997; Moutim *et al.*, 1999).

Foi demonstrado que o látex de *C. candamarcensis* é composto de cerca de 20% de carboidratos e 70% de proteínas (Baeza *et al.*, 1990), sendo que várias dessas proteínas são cisteíno-proteases, dotadas de atividade proteolítica de 5 a 8 vezes maior do que as encontradas em *C. papaya* (Baeza *et al.*, 1990; Gomes *et al.*, 2005; Teixeira *et al.*, 2008).

1.3 - As cisteíno-proteases

1.3.1 – Classificação e distribuição na natureza

As proteases, ou peptidases são enzimas que catalisam a hidrólise da ligação peptídica. Elas participam de vários processos fundamentais à vida como “turnover” de proteínas, resposta imune, sinalização celular, digestão, coagulação, apoptose, diferenciação celular, entre outros (Powers *et al.*, 2002).

As proteases podem ser divididas em exopeptidases e endopeptidases. As exopeptidases são proteases que clivam nas extremidades N e C-terminal, sendo denominadas amino-peptidases e carboxi-peptidases, respectivamente. As endopeptidases são capazes de clivar no interior da cadeia peptídica, sendo também chamadas de proteinases.

As endopeptidases podem ser divididas de acordo com o aminoácido que promove a catálise no sítio ativo, sendo os principais tipos: metalo, serino, aspartato e cisteíno-proteases (Rawlings & Barrett, 1993; Powers *et al.*, 2002). Há também as treonina e glutamato-proteases, além daquelas agrupadas por possuírem sítio ativo não-canônico ou desconhecido (Rawlings & Barrett, 1993; Rawlings *et al.*, 2008).

As cisteíno-proteases, também conhecidas como tiol-proteases, são enzimas que possuem em seu sítio ativo uma cisteína. Elas são amplamente distribuídas na

natureza sendo encontradas em diversos tipos de organismos, como mamíferos (Chapman *et al.*, 1997; Philchenkov, 2004), parasitas pluricelulares (Sajid & McKerrow, 2002, Robinson *et al.*, 2008; Shindo & Van der Hoorn, 2008) e unicelulares (Caffrey, 2001; Blackman, 2004; Mahmoudzadeh-Niknam *et al.*, 2004), em bactérias (Travis *et al.*, 1995), em vírus (Bazan & Fletterick, 1988; Krausslich & Wimme, 1988; Seipelt *et al.*, 1999) e em plantas (Mynott *et al.*, 1999; Azarkan *et al.*, 2003; Van der Hoorn, 2008; Gomes, 2008).

Segundo a base de dados MEROPS, que utiliza critérios evolutivos e estruturais para classificar as proteases (Barrett *et al.*, 2001; Rawlings *et al.*, 2008), em janeiro de 2009 estava registrada a ocorrência de 111.762 peptidases, sendo que dessas 18.997 eram cisteíno-proteases, divididas nos clãs CA, CD, CE, CF, CH, CL, CM, CN, CO e C-, que por sua vez são subdivididos em 59 famílias. O número de registro de cisteíno-proteases vem crescendo (Gomes, 2008), indicando que cada vez mais enzimas são descobertas e atribuídas novas funções a essa classe de proteínas.

1.3.2 – Características estruturais de cisteíno-proteases da família C1, clã CA

Dentro do clã com mais cisteíno-proteases, o clã CA, está a família C1, também conhecida como família da papaína. Essa família, detentora do maior número de peptidases do clã CA, possui como membros mais representativos as catepsinas e a própria papaína. Na família C1 são agrupadas as proteínas que possuem semelhança estrutural à papaína, a primeira protease a ter sua estrutura tridimensional desvendada (Drenth *et al.*, 1968; Wiederanders, 2003).

As cisteíno-proteases da família C1 são sintetizadas em forma de pré-proteínas. O prefixo “pré” refere-se à presença do peptídeo sinal, que pode ter entre cerca 10 a 25 aminoácidos, importante no endereçamento da protease para compartimentos intracelulares, como os lisossomos, ou para a sua excreção para o meio extracelular (Groves *et al.*, 1996; Kyndt *et al.*, 2007).

O prefixo “pro”, por sua vez, indica que a proteína é produzida na forma de zimogênio, ou seja, elas contêm uma pro-região, posterior ao pré-peptídeo, que também é conhecida como pro-domínio ou pro-peptídeo (Groves *et al.*, 1996; Wiederanders, 2003; Kyndt *et al.*, 2007). Na presença do pro-peptídeo a proteína se encontra no estado inativo e é conhecida como precursora. O tamanho do pro-

peptídeo varia entre 38 e 251 aminoácidos e são atribuídas diversas funções ao pro-peptídeo, dentre elas a inibição da enzima até que seja requisitada pela célula em sua forma madura. Essa inibição é proporcionada pelo impedimento estérico que a porção precursora causa no sítio ativo, uma vez que o pro-peptídeo se sobrepõe à fenda do sítio catalítico, impedindo a entrada do substrato (Groves *et al.*, 1996; Wiederanders, 2003). Os pro-peptídeos podem inibir especificamente suas cisteíno-proteases correspondentes, como demonstrado em estudos com proteínas recombinantes (Taylor *et al.*, 1995; Wiederanders, 2003). Também foi atestado que a pro-região é essencial para o envelhecimento correto de muitas cisteíno-proteases (Groves *et al.*, 1996; Bromme *et al.*, 2004).

A terceira região do gene das cisteíno-proteases corresponde à parte codificadora da proteína madura que tem tamanho médio de 220 aminoácidos (Wiederanders, 2003; Bromme *et al.*, 2004). Normalmente, as enzimas dessa família possuem sete resíduos de cisteína característicos, seis delas formando três pontes de dissulfeto localizadas entre os resíduos 22-63, 56-95 e 153-200 (numeração da cadeia da papaína), e uma cisteína 25 no sítio catalítico (Yamamoto *et al.*, 1991).

A estrutura terciária característica das moléculas do clã CA é composta por dois domínios, R e L, e entre esses domínios, o sítio ativo. A cisteína 25 está posicionada no domínio L, e a histidina 159 no domínio R, sendo altamente conservados em todas as cisteíno proteases. Também participa da catálise o resíduo da asparagina 175 (Grzonka *et al.*, 2001; Wiederanders, 2003).

Foi demonstrado que cinco subsítios são importantes na ligação com o substrato, sendo o S2, S1 e S1' os principais. Estes podem interagir com a cadeia principal e lateral dos substratos, enquanto que S3 e S2' podem interagir somente com a cadeia lateral dos aminoácidos (Turk *et al.*, 1998; Grzonka, 2001).

O mecanismo de catálise enzimática depende da tríade catalítica formada pela cisteína 25, histidina 159 e asparagina 175 e suas interações com o substrato. O grupo tiol da cisteína 25 localiza-se próximo do grupamento imidazol do resíduo de histidina 159. Este funciona como um doador e aceptor de prótons. Quando o grupamento imidazol tira um próton do grupamento sulfidrila (SH) da cadeia lateral da cisteína 25, este se torna mais nucleofílico (S⁻). A asparagina 175 é a responsável por estabilizar este estado (Sajid *et al.*, 2002). Assim, a formação do complexo enzima-substrato é seguido pelo ataque nucleofílico do S⁻ da enzima ao átomo de carbono da carbonila da ligação peptídica do substrato a ser hidrolizado,

ocorrendo a formação de um intermediário tetraédrico. Após a doação do próton pelo anel imidazólico da histidina 159, ocorre a liberação do fragmento C-terminal do substrato. A seguir, uma hidroxila de uma molécula de água reage com o intermediário, formando novamente o estado tetraédrico. O próton da molécula de água é recebido pelo anel imidazol da histidina. O estado tetraédrico é mais uma vez desfeito com a doação do próton pelo anel imidazol e o fragmento N-terminal do substrato é liberado, restaurando a enzima para mais um ciclo de reações (Storer *et al.*, 1994).

Algumas estruturas terciárias de cisteíno proteases já foram desvendadas como cruizipaina (Choe *et al.*, 2005), actinidina (Varughese *et al.*, 1992), catepsina K (Zhao *et al.*, 1997), catepsina B (Jia *et al.*, 1995), ervatamina B (Biswas *et al.*, 2003), entre muitas outras. Entre as proteínas da família Caricaceae que tiveram sua estrutura tridimensional elucidada tem-se como exemplo a papaína (Drenth, *et al.*, 1968; Yamamoto *et al.*, 1991), a quimopapaína (Maes *et al.*, 1996), a caricaína (Groves *et al.*, 1996) e glicil endopeptidase (O'Hara *et al.*, 1995), todas de *C. papaya*. Recentemente, Gavira *et al.* (2007) descreveram a estrutura da mexicanaína, uma cisteíno protease isolada de *Jacaratia mexicana*, outra planta da família Caricaceae e Gomes *et al.* (2008) elucidaram a estrutura da protease CMS1MS2, de *C. candamarcensis*.

1.3.3 – Cisteíno-proteases de *Carica candamarcensis* e suas propriedades biológicas

Assim como *C. papaya*, *Carica candamarcensis* é uma rica fonte de cisteíno-proteases. Várias proteínas dessa espécie já foram isoladas (Walraevens *et al.*, 1999; Gomes *et al.*, 2005; Teixeira *et al.*, 2008) mas, como outras proteases de plantas, ainda pouco se sabe sobre que funções essas peptidases exercem no organismo em que são encontradas (Wiederanders, 2003; Van der hoorn, 2008).

Dentre as proteínas isoladas do látex de *C. candamarcensis* estão as proteínas CC28, CC23, CC-Ia, CCI-b, CC-III, CC-IV (Gravina-de-Moraes *et al.*, 1994; Silva *et al.*, 1997; Walraevens *et al.*, 1999) e várias outras cisteíno-proteases que foram recentemente isoladas e caracterizadas quanto à sua sequência N-terminal (Gomes *et al.*, 2007, Salas *et al.*, 2008; Teixeira *et al.*, 2008). Enzimas muito

semelhantes às de *C. candamarcensis* também foram encontradas em *V. x heilbornii* e *V. stipulata*, sendo que algumas delas foram isoladas por métodos cromatográficos enquanto outras tiveram seus cDNAs sequenciados e a sequência de aminoácidos deduzida *in silico* (Kyndt *et al.*, 2007). Nestes estudos é evidente a produção de várias isoformas de cisteíno-proteases pelos membros do gênero *Vasconcellea*, apontando para uma sub-especialização das enzimas desse gênero (Pereira *et al.*, 2001; Kyndt *et al.*, 2007).

Recentemente o sequenciamento N-terminal de várias proteases de *C. candamarcensis* demonstrou a grande variedade de isoformas de cisteíno-proteases, que exibem atividades proteolíticas variadas e superiores às encontradas em *C. papaya*. Essas peptidases se originaram de cromatografias sequenciais em colunas de gel filtração e troca iônica, tendo o látex como material inicial (Teixeira *et al.*, 2008).

Algumas dessas proteínas já foram caracterizadas em relação à sua atividade proteolítica, sequência primária de aminoácidos e nucleotídeos. CC-III, por exemplo se destaca por ser a única cisteíno-protease de Caricaceae relatada com sítio de glicosilação. CC23, por sua vez, se destaca por ser a primeira enzima de *C. candamarcensis* com sua sequência nucleotídica desvendada (Pereira *et al.*, 2001).

No primeiro passo cromatográfico de gel filtração, duas frações protéicas do látex, denominadas P1G10 e P2G10 foram separadas. A fração P1G10 apresenta atividade cicatrizante sobre dermo-abrasões e úlceras gástricas na concentração de 0,1% e 10 mg/kg, respectivamente (Mello *et al.*, 2006; 2008). Também foi demonstrado que P1G10 possui baixa toxicidade aguda dérmica e não altera a atividade do citocromo P450 (Mello, 2005). Formulações de liberação controlada de P1G10 em β -ciclodextrina (Denadai *et al.*, 2006) também já foram obtidas. Devido a esses resultados que indicam P1G10 como um fármaco promissor na regeneração de lesões, estudos clínicos de fase II usando P1G10 como cicatrizante cutâneo estão sendo realizados. Entretanto, ainda pouco se conhece sobre como P1G10 estaria atuando na cicatrização nesses modelos.

A partir de P1G10, outras frações foram obtidas através de passagem em coluna de troca catiônica CM-Sephadex C-50, denominadas CMS1, CMS2 e CMS3. CMS1, uma fração com característica mais básica, apresenta atividade proteolítica maior em relação a CMS2, mais ácida (Teixeira *et al.*, 2008). CMS2 e CMS3 demonstraram atividade proliferativa sobre fibroblastos (Gomes, 2008) e CMS2

mostrou uma reação positiva contra anticorpos anti-EGF, demonstrando a presença de epítomos semelhantes a EGF nessa mistura (Gomes *et al.*, 2005).

As frações CMS1, CMS2 e CMS3 foram recromatografadas em coluna de troca catiônica Mono-S Sepharose, obtendo-se 14 proteínas das três frações. Das proteínas de CMS1, a proteína CMS1MS2 possui maior atividade proteolítica contra o substrato BAPNA dentre todas as proteínas derivadas da coluna CM-Sephadex (Teixeira *et al.*, 2008). Recentemente, a estrutura terciária dessa proteína foi elucidada após cristalização e difração de raios-X (Gomes *et al.*, 2008). Outra proteína importante obtida na cromatografia em Mono-S é CMS2MS2, que apresenta atividade mitogênica sobre fibroblastos e outras linhagens celulares, além de ter sido demonstrado que essa protease ativa a via das MAP-quinases (Gomes *et al.*, 2005; Gomes, 2008).

A estrutura primária de CMS2MS2 foi recentemente relatada por Gomes e colaboradores (2007), indicando ser uma proteína de 214 aminoácidos, com cerca de 23.590,76 Da, sem sítios de glicosilação e com diversos motivos estruturais primários e terciários comuns a cisteíno-proteases. Como relatado para outras proteínas de *Vasconcellea* (Walraevens *et al.*, 1999; Kyndt *et al.*, 2007), CMS2MS2 não possui um “loop” extra entre os resíduos 168-169 (numeração da cadeia da papaína), como ocorre com as cisteíno-proteases quimopapaína, caricaína e glicil-endopeptidase de *Carica papaya* (Kyndt *et al.*, 2007; Gomes *et al.*, 2007). A atividade proteolítica dessa enzima contra o substrato BAPNA é a maior entre as proteases da fração CMS2 e há uma preferência de clivagem em aminoácidos com características hidrofóbicas na posição P2 do substrato (Gomes, 2008).

A comparação e alinhamento da estrutura primária de CMS2MS2 com outras cisteíno-proteases de Caricaceae presentes no banco de dados do NCBI (disponível em: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) demonstrou semelhança de 83% com CC-III, 74% com quimopapaína, 68% com caricaína e 59% com papaína. Uma análise filogenética comprovou essa similaridade, agrupando CMS2MS2 com CC23, quimomexicaína e mexicaína, estas duas últimas proteases originárias de *Jacaratia mexicana*, também da família Caricaceae (Gavira *et al.*, 2007; Gomes *et al.*, 2007).

A função fisiológica de CMS2MS2 no látex e na formação do coágulo durante a injúria da planta ainda não é conhecida. Entretanto, outras funções foram descobertas para essa protease. Foi demonstrado que CMS2MS2 induz a proliferação de diversos tipos celulares, entre eles células de insulinooma (RINm5f),

osteoblastos de cultura primária e fibroblastos da linhagem L929. Essa ação foi mostrada no intervalo de 0,4 a 2 nM (Gomes *et al.*, 2005; Gomes, 2008). É possível que essa propriedade indutora de proliferação celular também seja a função que essa protease exerce durante o extravasamento do látex no local de injúria causado à planta, auxiliando no fechamento da abertura. Esse fato, entretanto, ainda não foi demonstrado.

Foi mostrado também que a atividade mitogênica de CMS2MS2 não é dependente de sua atividade proteolítica, pois a inibição da protease por E-64, um inibidor irreversível e específico de cisteíno-proteases, não resultou na perda dessa atividade biológica (Gomes, 2008).

Inibida ou não, a protease ativa a via das MAP-quinases, já que foi demonstrado que a fosforilação de ERK1/2 na presença de CMS2MS2 ocorre em níveis semelhantes ao controle positivo (estimulado por éster de forbol PMA e ionóforo de cálcio). A quantidade geral de ERK2 não-fosforilada, entretanto, parece estar inalterada durante a estimulação com CMS2MS2 (Gomes, 2008). Por apresentar atividade proliferativa, era esperado que CMS2MS2 apresentasse reação contra anti-EGF, como mostrado para a mistura de cisteíno-proteases derivadas de CMS2. Porém, Gomes e colaboradores (2005) demonstraram que CMS2MS2 e CMS2MS3 não apresentam reação com anti-EGF, enquanto que as três demais peptidases do pico CMS2 sim, indicando que os epítomos de EGF podem não estar presentes naquelas proteínas e, dessa forma, não serem responsáveis pela atividade proliferativa observada (Gomes *et al.*, 2005).

Vários estudos têm sido feitos a fim de esclarecer como CMS2MS2 induz a mitose, mas ainda não são conhecidos os mecanismos pelos quais a proteína poderia aumentar a proliferação celular e se essa capacidade também a torna apta para ser utilizada como fármaco, por exemplo, em processos de cicatrização dérmica e gástrica.

Devido às limitações e dificuldades de obtenção de CMS2MS2 por métodos cromatográficos, é necessária uma alternativa aprimorada de aquisição dessa proteína. Uma solução para esse problema seria a clonagem e expressão de CMS2MS2 em organismos heterólogos como *E. coli*, uma vez que a produção da proteína em bactéria poderia fornecer grandes quantidades dessa molécula.

1.4 – Amplificação, clonagem e expressão de genes de cisteíno-proteases de Caricaceae

Várias cisteíno-proteases da família Caricaceae tiveram seus genes amplificados por PCR e algumas delas foram clonadas e expressas em *E. coli*.

A primeira dessas peptidases a ser clonada e sequenciada foi a papaína, em 1986, por Cohen e seus colaboradores (Cohen *et al.*, 1986). Posteriormente, a papaína foi clonada e expressa em bactéria (Cohen *et al.*, 1990) e como pro-papaína foi expressa por Vernet e colaboradores (1991). Outras cisteíno-proteases de *C. papaya* também foram expressas e suas sequências gênicas conhecidas, como pro-caricaína (Revell *et al.*, 1993; Groves *et al.*, 1996), pro-quimopapaína (Taylor *et al.*, 1995) e glicil-endopeptidase em sua forma precursora (Baker *et al.*, 1996). Algumas dessas proteínas também foram cristalizadas e sua estrutura tridimensional resolvida por difração de raios X (Groves *et al.*, 1996, Maes *et al.*, 1996).

A expressão dessas proteínas e a visualização de sua estrutura terciária mostraram que para algumas proteínas, o pro-peptídeo é essencial para o enovelamento, provavelmente por diminuir a barreira energética necessária para o correto enovelamento da parte que corresponde à proteína madura (Bryan, 2002).

Atualmente vários estudos genéticos foram realizados com os membros do gênero *Vasconcellea*. Os trabalhos de amplificação de DNA genômico de diversas espécies de *Vasconcellea*, realizados por van Droogenbroeck e colaboradores (2004, 2006), e também por Kyndt e colaboradores (2005), demonstraram diferenças entre segmentos gênicos de cloroplasto, mitocondriais e também genômicos dessas espécies. Utilizando ferramentas de bioinformática e restrição enzimática dos fragmentos obtidos, esses trabalhos mostraram alto grau de hibridação entre as espécies, principalmente entre *C. candamarcensis*, *V. x heilbornii* e *V. stipulata* (van Droogenbroeck 2004, 2006; Kyndt *et al.*, 2005). Mais ainda, segundo os critérios utilizados por esses pesquisadores, o gênero *Vasconcellea* se encontra mais próximo de *Jacaratia* do que *Carica papaya* (Kyndt *et al.*, 2005), confirmando a proximidade filogenética já observada para a proteína CMS2MS2 (Gomes *et al.*, 2007).

Apesar da extensa análise de produtos de PCR, poucos genes do gênero *Vasconcellea* foram clonados e/ou expressos. As cisteíno-proteases de *V. stipulata* e *V. x heilbornii* foram amplificadas utilizando primers específicos para regiões internas

conservadas desse tipo de peptidase, obtendo-se seis cisteíno-proteases, quatro delas de *V. x heilbornii* (VXH-A, VXH-B, VXH-C e VXH-D) e duas de *V. stipulata* (VS-A e VS-B). Esses genes foram clonados para fins de sequenciamento, uma vez que se desejava obter as sequências pré-pro-peptídicas completas dessas proteínas (Kyndt *et al.*, 2007). Esse trabalho compara a sequência nucleotídica de todas as cisteíno-proteases até então registradas, mais uma vez mostrando a semelhança entre as cisteíno-proteases de Caricaceae, ao mesmo tempo em que apresenta mais uma vez várias isoformas dessas peptidases nas espécies de *Vasconcellea*.

A clonagem de proteínas de *C. candamarcensis* também já foi realizada. Pereira e colaboradores (2001) realizaram a clonagem do possível gene que codifica a proteína CC23 (número de acesso no NCBI: AY035986). Essa proteína apresenta alta homologia com VS-B, VXH-C, CC-III, mexicaína, quimomexicaína e as cinco isoformas de quimopapaína.

Recentemente, tentativas de clonagem da proteína CMS2MS2 foram realizadas (Junqueira, 2007). Nesse trabalho, duas sequências de cisteíno-proteases foram obtidas: a porção precursora de uma cisteíno-protease (AF417109) e uma sequência de cisteíno-protease (AY194124). Entretanto, a expressão dessa última em vetor pMal -p2 não foi bem sucedida.

Dentre essas, CC-23, amplificada a partir de DNA genômico, possui a maior homologia com a sequência de aminoácidos de CMS2MS2 (94%), seguido da sequência de cisteíno-protease AY194124 (61%), embora estas sejam sequências incompletas que possuem apenas até o resíduo de glicina 176, dos 214 aminoácidos de CMS2MS2. A sequência precursora AF417109 possui fraca homologia com CMS2MS2, apesar de apresentar alta semelhança com proteínas precursoras de outras plantas, utilizando a ferramenta “blastp” do NCBI (disponível em: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

1.5 - O ciclo celular em eucariotos

Devido ao fato da protease CMS2MS2 apresentar atividade mitogênica, torna-se pertinente uma breve descrição do ciclo celular, processo dependente de fatores de crescimento.

A divisão celular abrange um conjunto de eventos que ocorrem no citoplasma e núcleo das células, caracterizados por fases que se efetuam de forma sequenciada, incluindo a intérfase, prófase, metáfase, anáfase e telófase, culminando na produção de células-filhas (Mitchison, 1988). Essa divisão resulta na duplicação do número de células e, posteriormente, na transcrição e tradução de genes que codificam proteínas constitutivas do fenótipo característico de cada linhagem celular (Murray & Szostak, 1985; Oshima 1991).

O período que compreende do início até o final da mitose é chamado ciclo celular, durando de 18 a 24 horas na maioria das células somáticas animais (Mitchison, 1988). A intérfase ocupa a maior parte do ciclo celular, e compreende as fases G_1 , S e G_2 . Os eventos das fases G_1 e G_2 são regulados por “checkpoints”, etapas em que algumas proteínas reguladoras controlam o prosseguimento das fases (Pardee, 2002).

A fase G_1 dura em média de 6 a 12 horas, sendo o período de maior duração dentro do ciclo celular. Para começar o ciclo, a célula em fase G_1 precisa ser estimulada por fatores de crescimento (Harding *et al.*, 2003). Durante a fase G_1 , a célula passa por vários “checkpoints”, controlados principalmente pela proteína *p53* e pelas ciclinas do tipo D (Pardee, 2002). A proteína *p53* tem a capacidade de alongar o tempo da fase G_1 a fim de possibilitar o reparo de erros na sequência de DNA (Hupp *et al.*, 2000), enquanto as ciclinas do tipo D irão regular a sinalização induzida pelos fatores de crescimento (Pardee, 2002).

O ponto da fase G_1 em que o comprometimento com a duplicação ocorre e a célula não precisa mais de fatores de crescimento para completar o ciclo celular é chamado de ponto de restrição. Esse ponto é atingido com o acúmulo de uma proteína lábil denominada proteína R, cuja síntese é sensível ao estímulo por fatores de crescimento (Pardee, 2002).

Se as condições não são favoráveis, a célula não atinge o ponto de restrição, sendo impedida de prosseguir para a fase S (Weinert & Hartwell, 1988; Nurse, 1990; Oshima, 1991). As células nessa situação podem ficar restritas em um estado de quiescência denominado G_0 . Muitas células podem ser estimuladas a saírem desse estado e entrarem no ciclo celular por meio de estímulos proliferativos provocados por fatores de crescimento (Oshima, 1991; Harding *et al.*, 2003).

Transcorrida a fase G_1 , a célula entra na fase S. Nessa fase o genoma é duplicado num período médio de 6 a 8 horas (Nurse, 1990). Uma vez começada a

duplicação do material genético, a célula está comprometida a terminar a replicação do DNA. A fase S, assim como G_2 e M já não respondem aos fatores de crescimento (Oshima 1991; Pardee 2002).

Em seguida ocorre a fase G_2 , uma preparação para a organização e separação do material genético durante a mitose (Mitchison, 1988; Oshima, 1991). Nessa fase, os checkpoints evitam a segregação de cromossomos defeituosos (Pardee, 2002).

A fase M, que é a mitose propriamente dita, é a mais breve das etapas do ciclo celular, tendo aproximadamente uma hora de duração. Nessa etapa o material genético, condensado em cromossomos, se organiza em fusos e é separado entre as células filhas, que são formadas assim que o citoplasma se rompe e se reorganiza na citocinese (Murray & Szostak, 1985; Mitchison, 1988; Oshima, 1991).

1.6 - Fatores de crescimento celular

Existem dois tipos principais de fatores de crescimento: os fatores de competência e de progressão.

Os fatores de competência são de expressão genômica imediata ou primária, capazes de estimular o início do processo mitótico quando as células, por carência nutricional, se encontram no estado de quiescência ou G_0 . Eles são capazes de induzir a expressão de genes envolvidos no processo de divisão celular, sendo, portanto, empregados como complementos do meio de cultura para retirar a célula do estado quiescente e devolvê-la ao ciclo celular (Pardee, 1989) O Fator de Crescimento Derivado de Plaquetas (PDGF; "Platelet-Derived Growth Factor") é um exemplo de fator de competência (Pardee, 1989). Outros exemplos de fatores de competência são FGF ("fibroblast growth factor") e bombesina (Pardee, 1989; Rozengurt, 1986).

Os fatores de progressão, por sua vez, são capazes de estimular a célula a entrar na fase S, quando esta já foi estimulada previamente por um fator de competência e tenha sido retirada do estado de quiescência para a fase G_1 . Portanto, os fatores de progressão não são capazes de retirar as células do estado G_0 , e só produzirão estímulo mitogênico se a célula estiver em G_1 . Isto pode ser observado em células epiteliais que necessitam ser tratadas com fatores de

competência antes do tratamento com insulina, que só possui efeito mitogênico nestas condições (Ross & Ballard, 1988). As somatomedinas também são consideradas fatores de progressão. As somatomedinas do tipo A e C, por exemplo, são tão potentes quanto o MSA ('multiplication-stimulating activity'), também da família das somatomedinas (Pierson & Temin, 1972; Stiles *et al.*, 1979).

Com relação ao mecanismo de ação, os fatores de crescimento interagem com receptores proteicos transmembrânicos específicos que possuem regiões conservadas tanto em domínios citoplasmáticos quanto em domínios extracelulares (Ward *et al.*, 1995). Estes receptores estão presentes em células normais e transformadas (Wahl & Carpenter, 1987). Os receptores de fatores de crescimento são proteínas tirosina-quinases, tendo um sítio de ligação no domínio extracelular e atividade quinase em seus domínios citoplasmáticos (Iwasaki *et al.*, 1998).

A interação entre o fator de crescimento e seu receptor é bem exemplificada pelo Fator de Crescimento Epidermal (EGF; "Epidermal Growth Factor"), um dos fatores de crescimento mais bem estudados e um fator de progressão para fibroblastos (Leof *et al.*, 1982). O processo de ativação mitogênica é desencadeado após a ligação de EGF ao seu receptor (Carpenter & Cohen, 1990).

O EGF humano é uma proteína de 53 aminoácidos e como a maioria dos outros fatores de crescimento, é sintetizada na forma de precursor, com unidades repetidas da molécula de EGF (Massagué & Pandiella, 1993). Já o receptor de EGF humano possui 1186 aminoácidos e consiste de uma glicoproteína transmembrânica (Ullrich *et al.*, 1984). A ligação de EGF ao seu receptor ocorre na região extracelular, a qual é dividida em quatro domínios, I-IV (Bajaj *et al.*, 1987; Ward *et al.*, 1995).

A ligação do EGF ao seu receptor entre os domínios extracelulares I e III, faz com que este mude sua conformação tridimensional, tornando-o capaz de interagir com outro receptor que também tenha se ligado ao EGF. Portanto, é formado um complexo de proporção EGF:receptor de 2:2 (Ogiso *et al.*, 2002). Essa dimerização na superfície extracelular aproxima os domínios citoplasmáticos que possuem atividade tirosina-quinase, perto o suficiente para se autofosforilarem. A fosforilação do receptor de EGF, assim como ocorre com outros receptores de fatores de crescimento com domínio tirosina-quinase, ativa várias vias de sinalização, como a das MAP-quinases, no citoplasma da célula, o que posteriormente estimula a transcrição de genes envolvidos na proliferação (Carpenter & Cohen, 1990).

1.7 - Proteases como fatores de proliferação

Muitas proteases com capacidade de induzir a proliferação celular já foram descritas (Zetter *et al.*, 1976; Moses, 1997; DeClerck *et al.*, 2004), podendo agir como fatores de proliferação de várias maneiras.

As proteases podem clivar proteínas da superfície celular importantes por mediarem o controle do crescimento por contato com células vizinhas (Zetter *et al.*, 1976). Elas também podem clivar proteínas da matriz extracelular liberando fatores de crescimento que estavam alocados neste emaranhado protéico (Moses, 1997; DeClerck *et al.*, 2004). Alternativamente, as proteases podem clivar e solubilizar fatores de crescimento os quais estão complexados com proteínas da matriz extracelular (DeClerck *et al.*, 2004). Tais fatores, então, ficam livres para interagir com as células próximas, promovendo proliferação e migração celular (Bauvois, 2004).

Uma outra forma de ativação da mitogênese foi descrita para o Fator de Crescimento do Plerocercóide (PGF: “plerocercoid growth factor”), isolado do verme parasita *Spirometra mansonioides*. Esse fator é capaz de acelerar o crescimento de células do hospedeiro de uma maneira similar ao hormônio do crescimento. O PGF é uma cisteíno-protease com aproximadamente 50% de identidade com outras cisteíno-proteases de mamífero e nenhuma homologia com o hormônio do crescimento humano (Phares, 1996). Porém, PGF é capaz de interagir com receptores do hormônio de crescimento e provocar efeitos biológicos semelhantes aos mediados por estes hormônios de mamíferos (Phares, 1996).

Além disso, a proliferação celular pode ser induzida por enzimas proteolíticas através de receptores ativados por proteases (PAR: “protease activated receptor”), que são receptores acoplados à proteína G (GPCR: “G-protein coupled receptor”). Até o momento, quatro tipos de PAR foram identificados (Ossovskaya & Bunnnett, 2004).

PAR-1, por exemplo, possui 425 aminoácidos com sete domínios transmembrânicos hidrofóbicos típicos de GPCR. O domínio amino-terminal está localizado extracelularmente, e possui um sítio de clivagem para a serino-protease trombina (Vu *et al.*, 1991). Além de participar do processo de coagulação, foi demonstrado que a trombina também é capaz de ativar mitogenicamente células musculares lisas de vasos sanguíneos (Antonaccio *et al.*, 1993), células endoteliais

(Brass & Molino, 1997) e fibroblastos (Chen *et al.*, 1994), sendo importante no mecanismo de cicatrização e reparo de tecidos lesados (Déry *et al.*, 1998).

Durante a ativação do receptor PAR-1, ocorre a clivagem proteolítica da porção N-terminal do receptor pela trombina. Após essa clivagem, portanto, é formada uma nova extremidade amino-terminal que funciona como ligante para o segundo “loop” extracelular do receptor PAR-1 (Ossovskaia & Bunnett, 2004). Essa ligação ativa a proteína G complexada ao PAR-1, disparando várias vias de sinalização no citoplasma da célula, podendo culminar em processos como diferenciação ou mitose celular.

Algumas proteínas de origem vegetal capazes de desencadear respostas proliferativas foram descritas atuando em diversos sistemas biológicos.

A fitohemaglutinina, por exemplo, é uma lecitina isolada de *Phaseolus vulgaris* e foi uma das primeiras substâncias descritas com capacidade de induzir mitose celular em linfócitos humanos, tendo sido empregada para tal fim desde então (Obaru, 1986).

Das sementes de *Dioclea lehmanni* também foram isoladas duas lecitinas denominadas DI-1 e DI-2, com propriedades mitogênicas. A primeira possui forma semelhante a Concanavalina A, outra lecitina vegetal capaz de desencadear mitose em linfócitos PBL (“Peripheral Blood Lymphocyte”). A segunda é possuidora de estrutura dimérica, capaz de induzir a proliferação de células PBL e eritroaglutinação de hemácias humanas (Pérez, 1998).

Nosso grupo tem mostrado que as proteases do látex de *Carica candamarcensis* têm capacidade de estimular a proliferação de células de várias linhagens celulares de mamíferos. Devido ao potencial uso das proteases de *C. candamarcensis* como fármacos, a elucidação dos mecanismos envolvidos durante a indução da mitose por essas proteases é de essencial importância. Nesse contexto, CMS2MS2 se destaca, por ser uma protease isolada do látex que possui atividade proliferativa comprovada sobre diversas linhagens celulares e cuja ação é dependente de MAP-quinase. Entretanto, esses conhecimentos não são suficientes para esclarecer como essa e outras proteases pertencentes à fração P1G10 estão participando de processos proliferativos e/ou migratórios observados durante a cicatrização induzida por P1G10.

A obtenção de CMS2MS2 por processos cromatográficos a partir do látex de *Carica candamarcensis* é um processo trabalhoso e demorado. Diante do exposto, o

objetivo desse trabalho foi produzir CMS2MS2 recombinante em sistema heterólogo como *Escherichia coli*, com o propósito de fornecer maiores quantidades de proteína para posteriores estudos da atividade biológica e caracterização estrutural.

Objetivos

Gerais

Clonar, expressar e purificar a cisteíno-proteinase mitogênica CMS2MS2 de *C. candamarcensis* a fim de verificar sua atividade biológica proliferativa.

Específicos

- Amplificar o cDNA da proteína CMS2MS2 por PCR e clonagem do amplicon no vetor pCR 2.1.
- Verificar, por sequenciamento, a presença do gene da proteína CMS2MS2 em vetor pCR 2.1
- Sub-clonar o inserto contendo o gene de CMS2MS2 no vetor de expressão pET28a(+)-TEV, que deve produzir a proteína recombinante fusionada a cauda de histidina.
- Expressar e purificar a proteína recombinante em cromatografia de afinidade ao níquel a fim de produzir grandes quantidades de proteína.
- Avaliar a atividade biológica da proteína recombinante.

3- Materiais e Métodos

3.1- Materiais

3.1.1- Reagentes Químicos

Ácido acético glacial P.A., bicarbonato de sódio, cloreto de sódio, cloreto de potássio, clorofórmio, formaldeído, fenol P.A., hidróxido de sódio, glicina e tolueno P.A. (Labsynth Produtos para Laboratório Ltda, Diadema, SP, Brasil).

Acetato de sódio P.A., acetato de potássio P.A., etanol absoluto, etileno-diamino-tetracetato dissódico (EDTA) P.A., glicerol, fosfato monobásico de potássio P.A., fosfato dibásico de potássio trihidratado, fosfato monobásico e dibásico de sódio P.A., triton X-100 e metanol absoluto (Merck, Darmstadt, Alemanha).

Acrilamida (Pharmacia LKB, Uppsala, Suécia).

Albumina de ovo ou ovalbumina (OVA), cloreto de magnésio, cloreto de potássio, dimetil sulfóxido (DMSO), hidróxido de potássio, lisozima, persulfato de amônio (APS), Tris base, Ampicilina, RNase A, Dietil pirocarbonato (DEPC) brometo de etídeo, ácido 3-(N – morfolino) – propasulfônico (MOPS) glicerol, glicose, triptona e extrato de levedura, X-gal (5-bromo-4-cloro-3-indoxil-b-D-galactosídeo), IPTG (isopropil b-D-tiogalactopiranosídeo), estreptomicina, ácido N-(2-hidroxi-2-etil)-piperazino-N'-etaneosulfônico (Hepes), gentamicina, pirofosfato de sódio, 2,5-difenil-oxazol (PPO), 2,2-*para*-fenileno-bis-(5-fenil-oxazol) (POPOP) (Sigma Chemical Co, St. Louis, MD, EUA).

Anfotericina B (Bristol Meyers Squibb Indústria Farmacêutica S/A, São Paulo, SP, Brasil).

Ácido clorídrico P.A. (Vetec Química Fina Ltda, Rio de Janeiro, RJ, Brasil).

Agar (Isofar, Petrópolis, RJ, Brasil).

Agarose (Hispanagar, Burgos, Espanha).

Taq polimerase (Phoneutria, Belo Horizonte, MG, Brasil).

Taq polimerase, Taq platinum, dNTP, cloreto de magnésio, tampão Taq 10X, “primers” CMS2MS2 Forward e Reverse, “primers” universais M13 “Forward” e “Reverse”, T7 “promoter” e “terminator”, plasmídeo pCR 2.1, T4 DNA ligase, tampão

ligase 10X, λ DNA/ HindIII fragments, 100pb DNA ladder e 1Kb DNA ladder (Invitrogen, Carlsbad, Calif3rnia, EUA).

Plasm3deo pET28a(+)-TEV, fornecido pelo professor Nilson Zanchin do LNLS, Campinas, S3o Paulo, Brasil.

Enzimas de restri3o3o EcoRI, HindIII, Xho I e Sall, albumina bovina (BSA), ur3e3ia (GE Healthcare Life Sciences, Chalfont St. Giles, UK)

Azul de bromofenol, bis-acrilamida, corante azul-brilhante de coomassie G-250, dodecil sulfato de s3odio (SDS) e N',N',N',N'-tetrametil-etileno-diamina (TEMED) (Bio-Rad Laboratories, Hercules, CA, EUA).

[H³-metil]-timidina em solu3o3o aquosa (GE HealthCare, Buckinghamshire, Inglaterra).

Meio de cultura RPMI 1640 desidratado, solu3o3o de 5-bromo-4-cloro-3-idolil fosfato (BCIP) e solu3o3o de nitro azul de tetraz3olio (NBT) (Gibco-BRL, Gaithersburg, MD, EUA).

Nitrog3enio l3iquido (White Martins do Brasil S/A, Contagem, MG, Brasil).

3gua deionizada e destilada fornecida pelo Laborat3orio de Biologia Molecular de Produtos Naturais (ICB, UFMG, Belo Horizonte, MG, Brasil).

3gua milli-Q fornecida pelo Laborat3orio de Subst3ancias Antitumorais (ICB, UFMG, Belo Horizonte, MG, Brasil).

3.1.2- Solu3o3es

Tamp3o "grinding"

	mM
Tris-base	180
Cloreto de l3itio (LiCl)	90
EDTA	4.5
SDS 1 %	
Ajustar pH para 8.2 com HCl.	

TLE

	mM
Tris-base	200
Cloreto de lítio (LiCl)	100
EDTA	5
Ajustar pH para 8.2 com HCL	

Fenol equilibrado com TLE

Extrair o fenol com um volume igual de solução TLE, mais 0,5 mL de 15 m NaOH, trazendo o pH para perto de 8. Extrair mais duas vezes com TLE.

10 X MOPS "running buffer"

	mM
MOPS, pH 7	400
Acetato de sódio	100
EDTA	10

Meio Luria-Bertani (LB) Líquido

	g/L
Triptona (peptona).....	10
NaCl.....	10
Extrato de levedura	5

Mexer até dissolver. Ajustar o pH para 7,0 com NaOH 5N. Ajustar o volume da solução para 1L. Esterilizar em autoclave 20 minutos, 120 °C.

Meio LB Sólido

	g/L
Triptona (peptona).....	10
Extrato de levedura	5
NaCl	5
Agar	15

Mexer até dissolver. Ajustar o pH para 7,5 com NaOH 5N. Ajustar o volume da solução para 1L. Esterilizar em autoclave 20 minutos, 120 °C.

Tampão TAE (Tris Acetato EDTA) 50X

Tris-acetato 2M

EDTA 0,05M

Tampão TE (Tris EDTA)

mM

Tris-HCl pH 8,0 10,0

EDTA 0,1

Gel de Agarose 1%

Agarose 0,30 g

H₂O deionizada 29,00 ml

TAE 50X 0,60 ml

Brometo de etídeo (para corar géis de agarose)

10 µl de brometo de etídeo 10 mg/ml

Água destilada q.s.p. : 100 ml

Tfbl (para preparo de células quimiocompetentes)

mM

Acetato de potássio..... 30

Cloreto de manganês 50

Cloreto de rubídio 100

Glicerol 15 %

Ajustar pH para 5.8.

Tfb II (para preparo de células quimiocompetentes)

mM

Na-MOPS 10

Cloreto de cálcio..... 75

Cloreto de rubídio 10

Glicerol 15 %

Ajustar pH para 7.0.

Tampão de Amostra para eletroforese em gel de agarose (SDS Buffer 6X)

Azul de bromofenol	0,25% p/v
Xileno cianol	0,25% p/v
Glicerol	0,25% v/v

Solução de “Ressuspensão” para “mini-prep”

	mM
Glicose	50
EDTA pH 8,0.....	10
Tris-HCl pH 8,0	25

Solução de “Lise” para “mini-prep”

NaOH 0,2 M
SDS 1%

Solução de “Neutralização” para “mini-prep”

Acetato de Potássio 5 M pH 4,8.

Tampão EcoRI

	mM
Tris.HCl	100
Cloreto de magnésio (MgCl ₂)	5
Cloreto de sódio (NaCl)	50
Bovine serum albumin (BSA) 100 µg/ml (pH 7,5)	

Tampão K (Hind III e Xho I)

	mM
Tris.HCl	200
Cloreto de magnésio (MgCl ₂)	100
Ditiotreitol (DTT)	10
Cloreto de potássio (KCl)	1000

Tampão M (Sal I)

	mM
Tris.HCl	100
Cloreto de magnésio (MgCl ₂)	100
Ditiotreitol (DTT)	10
NaCl	500

Tampão de ligação para cromatografia de afinidade ao níquel

	mM
Fosfato de sódio	20
NaCl	500
Imidazol	25
Ajustar pH para 7.4	

Tampão de eluição para cromatografia de afinidade ao níquel

	mM
Fosfato de sódio	20
NaCl	500
Imidazol	250
Ajustar pH para 7.4	

Solução para o gel separador desnaturante 12% acrilamida (Laemmli, 1970)

Solução de acrilamida:bis-acrilamida 29:1 (p/p) 30% (p/v)	1,2 ml
Tampão Tris-HCl 1,0 mM pH 8,8; 0,4 % SDS (p/v)	0,75 ml
Água deionizada	1,02 ml
Solução aquosa de persulfato de amônio 10% (p/v)	25 µl
TEMED	5 µl

Solução para o gel concentrador desnaturante 5% acrilamida (Laemmli, 1970)

Solução de acrilamida:bis-acrilamida 29:1 (p/p) 30% (p/v)	166,8 µl
Tampão Tris-HCl 1,0 mM pH 6,8; 0,4 % SDS (p/v)	250,0 µl
Água deionizada	571,2 µl
Solução aquosa de persulfato de amônio 10% (p/v)	10 µl

TEMED 2 µl

Solução fixadora para géis

Metanol P.A. 50 ml

Ácido acético glacial P.A. 10 ml

Água destilada q.s.p. 100 ml

Solução de azul-brilhante de coomassie para coloração de géis

Corante azul-brilhante de Coomassie G-250 125 mg

Ácido acético glacial P.A. 10 ml

Água destilada q.s.p. 100 ml

Solução descorante de géis

Metanol P.A. 5 ml

Ácido acético glacial P.A. 7 ml

Água destilada q.s.p. 100 ml

Solução tampão de amostra eletroforese desnaturante

Tampão Tris-HCl 1,0 mM pH 6,8; 0,4 % SDS (p/v) 0,7 ml

Glicerol 0,3 ml

SDS 0,1 g

Azul de bromofenol 0,12 mg

Tampão de corrida eletroforese desnaturante

Tris base 0,755 g

Glicina 3,6 g

SDS 0,25 g

Água destilada q.s.p. 250 ml

Tampão PBS 1X sem EDTA pH 7,4

Fosfato dibásico de sódio P.A. 1,15 g

Fosfato monobásico de potássio P.A. 0,25 g

Cloreto de sódio 8,20 g

Cloreto de potássio 0,20 g

Água destilada q.s.p. 1 L

O pH é ajustado para 7,4 com solução de ácido fosfórico 0,1 M e o tampão é esterilizado por autoclavação a 120°C durante 20 minutos.

Tampão PBS/EDTA pH 7,4

Fosfato dibásico de sódio P.A. 1,15 g

Fosfato monobásico de potássio P.A. 0,25 g

Cloreto de sódio 8,20 g

Cloreto de potássio 0,20 g

EDTA 1,37 g

Água destilada q.s.p. 1 L

O pH é ajustado para 7,4 com solução de ácido fosfórico 0,1 M e o tampão é esterilizado por autoclavação a 120°C durante 20 minutos.

Meio RPMI 1640 pH 7,4

Meio RPMI 1640 desidratado 10,43 g

Bicarbonato de sódio 1,20 g

Solução aquosa de Hepes 1,0 M pH 7,4 10 µl

Ampicilina sódica 0,10 g

Estreptomicina 0,10 g

Gentamicina 0,05 g

Anfotericina B 0,05 g

Água deionizada q.s.p. 1,00 L

Esterilizado por filtração em membrana com porosidade absoluta de 0,22 µm.

Solução de aquassol para leitura de cintilação

PPO 4 g

POPOP 0,2 g

Triton X-100 333,0 ml

Tolueno q.s.p. 1,00 L

3.1.3- Biológicos

Cepas de *E. coli* TOP10 e BL21-DE3 cedidas pelo professor Alfredo Miranda Góes, Departamento de Bioquímica e Imunologia da Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil.

Fibroblastos tumorigênicos de camundongo (linhagem L929), gentilmente cedido pelo Dr. Hugo Armelin, Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo (USP), São Paulo, SP, Brasil.

Cultura primária de osteoblastos extraída da calvária de ratos wistar neonatos cedidas pela Dr^a. Míriam Teresa Paz Lopes, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil.

Soro fetal bovino (FBS) (Amershan Life Science, Buckinghamshire, Inglaterra).

Solução aquosa de tripsina (ATV) (Instituto Adolpho Lutz, São Paulo, SP, Brasil).

3.2 - Métodos

3.2.1 - Extração de RNA

O material utilizado para extração de RNA foram folhas da planta *C. candamarcensis*, coletadas em La Serena, Chile, que foram acondicionadas em gelo seco assim que destacadas da planta. O RNA foi extraído utilizando o método de extração fenólica (Ausubel *et al.*, 1987).

Inicialmente, as folhas foram trituradas em nitrogênio líquido com ajuda de um homogeneizador, até que se tornasse um pó bastante fino. Imediatamente após, o pó de folhas era transferido a um recipiente com 150 ml de tampão "grinding" além de 50 ml de fenol equilibrado com TLE e misturado por 2 minutos, com agitação moderada. Após homogeneização 50 ml de clorofórmio foram adicionados e a mistura foi novamente homogeneizada e incubada a 50 °C por 20 minutos. Em

seguida, a solução foi centrifugada a 10000 rpm, por 20 minutos, a 4 °C, para extração da parte aquosa.

A etapa de extração com fenol-TLE e clorofórmio foi repetida três vezes e posteriormente realizada uma última extração apenas com 50 ml de clorofórmio. Ao produto final da extração foi adicionado 8M de LiCl (1/3 do volume) e o RNA foi precipitado a 4°C, overnight.

No dia seguinte, a solução foi centrifugada a 10000 rpm, 4°C, por 20 minutos para a recuperação do RNA precipitado. Este foi lavado com 2M LiCl e centrifugado novamente. O precipitado foi ressuspenso em 5 ml de água livre de RNase (água DEPC) e novamente adicionado 8M LiCl (1/3 do volume). Essa mistura foi incubada durante 3 horas, novamente centrifugada e lavada com 2M LiCl.

O RNA foi então ressuspenso em 2 ml de água DEPC, 200 µl de acetato de sódio, 5.5 ml de etanol absoluto e precipitado em banho de gelo seco/etanol por 30 minutos. Depois dessa etapa, centrifugou-se a 10000 rpm, 4°C, por 15 minutos. O RNA foi ressuspenso em 20 µl de água livre de RNase.

A concentração de RNA total foi estimada por espectrofotometria a A_{280}/A_{260} e por análise de eletroforese em gel de agarose/formaldeído (Ausubel *et al.*, 1987).

O gel consiste de uma solução de agarose 1 %, 10 % do tampão MOPS 10 X e 18 % de formaldeído 12,3 M. As amostras de RNA foram preparadas em 10% de tampão MOPS 10 X, 18 % de formaldeído a 12,3 M e 50 % formamida deionizada. Em seguida, elas foram aquecidas a 55°C por 15 minutos. O tampão de corrida foi acrescentado após o aquecimento.

A eletroforese foi desenvolvida horizontalmente a 5 V/cm. Após a corrida, o gel foi lavado duas vezes com acetato de amônio a 0,5 M e corado com uma solução de brometo de etídeo 0,5 µg/ml e 0,5 M de acetato de amônio durante 40 minutos.

O gel foi observado em transiluminador ultravioleta e documentado utilizando a câmera fotográfica Kodak DC40.

Em seguida, o RNA foi aliquoteado em quatro tubos do tipo “eppendorf”, com cerca de 500 µg de RNA cada. Essas amostras foram secas no speed vac e armazenadas a -70°C até sua utilização.

3.2.2- Síntese de cDNA

A síntese de cDNA foi feita utilizando First-Strand cDNA Synthesis Kit (Amersham Biosciences/GE HealthCare Lifesciences). Nesse processo ocorre a síntese de RNA complementar pela reação RT-PCR (“reverse transcription-polymerase chain reaction”).

O RNA foi ressuspendido primeiramente em 20 µl de água livre de RNase. Uma alíquota de 1 µl de RNA foi colocada em um tubo “eppendorf” juntamente com 5,5 µl de água. Essa mistura foi aquecida a 65°C por 10 minutos. Antes da próxima etapa, o RNA dissolvido era incubado durante 5 minutos no gelo.

Depois eram adicionados 1 µl de solução de DTT, 1 µl do primer NotI-d(T)₁₈, 5 µl do “mix” para a reação, completando para 15 µl de volume final com água livre de RNase.

A reação foi, então, incubada por uma hora, em banho-maria a 37°C. Logo em seguida a transcriptase reversa foi desnaturada a 90°C por 5 minutos.

3.3.3 - Amplificação do gene de CMS2MS2 a partir de cDNA

A proteína CMS2MS2, purificada do látex de *C. candamarcencis* foi totalmente sequenciada pelo método de degradação de Edman por Gomes *et al.* (2007). A sequência obtida foi a seguinte:

YPESIDWRQKGAVTPVKDQNPCGSCWAFSTVATVEGINKIVTGKLISLSEQELLDCD
RRSHGCKGGYQTTSLQYVVDNGVHTEKVYPYEKKQKCRADKQGPWVKITGYK
RVPSNDEISLIKAIATQPVSVLVESKGRAFQFYKGGVFGGPCGTKLDHAVTAVGYGK
DYILIKNSWGLRWGDKGYIKIKNASGNSEGICGVYKSSYFPIKGYQ

Para amplificar o gene da protease CMS2MS2 foi desenhado um par de “primers” degenerados a partir da sequência de aminoácidos:

CMS2MS2 Foward: 5' TAY CCN GAR AGY ATH GAY TGG 3', correspondente à região N-terminal da proteína.

CMS2MS2 Reverse: 5' YTG RTA NCC YTT DAT NGG RAA RTA 3', correspondente à região C-terminal da proteína.

A mistura da reação de PCR era composta de: 1U de *Taq* DNA polimerase, 5 mM dNTP, 5 µl tampão *Taq* 10X, 50 pmoles de cada primer, 2,5 µl de cDNA, q.s.p. 50 µl de água livre de DNase. O controle negativo era constituído pela mesma mistura, com exceção do DNA. O programa utilizado para amplificação foi:

- 1) 94°C, 2 minutos – desnaturação do DNA
- 2) 39°C, 1 minuto – anelamento dos primers
- 3) 72°C, 2 minutos – extensão e amplificação do DNA
- 4) 95°C, 1 minuto
- 5) 39°C, 1 minuto
- 6) 72°C, 2 minutos
- 7) Repetição de 4 a 6, 34 vezes
- 8) 72°C, 15 minutos – extensão final
- 9) 4°C infinito

3.3.4 - Análise e quantificação do produto de PCR

O produto de PCR foi analisado por eletroforese de ácidos nucléicos em gel de agarose 1%, em TAE 1X e corado com solução de brometo de etídeo. A amostra foi diluída em tampão de amostra de DNA 6X (1/6 do volume). A eletroforese foi realizada horizontalmente, sob corrente constante de 80 V, 60 minutos. O gel foi visualizado sob luz ultravioleta, e as bandas de DNA referentes à amplificação eram comparadas com os padrões de peso molecular 100 bp DNA ladder (Invitrogen), 1 kb DNA Plus ladder (Invitrogen) e/ou λ DNA Hind III (Invitrogen). Para documentação foi utilizada câmara fotográfica Kodak modelo DC40 e a análise das bandas realizada com o software Kodak versão 1D-3.5. As fotos dos géis também foram documentadas no equipamento ImageMaster VDS (Pharmacia Biotech).

3.3.5 - Ligaç o em pCR 2.1

O vetor pCR 2.1 (TA Cloning Kit, Invitrogen), representado na figura 1, foi utilizado para clonagem de produtos de PCR amplificados utilizando a enzima *Taq* DNA polimerase, que adiciona uma adenina na extremidade 3' do fragmento amplificado por PCR. O vetor possui em suas pontas 3' uma mol cula de timidina que, durante a ligaç o por T4 DNA ligase, se unir o  s adeninas livres do produto de PCR por pontes de hidrog nio.

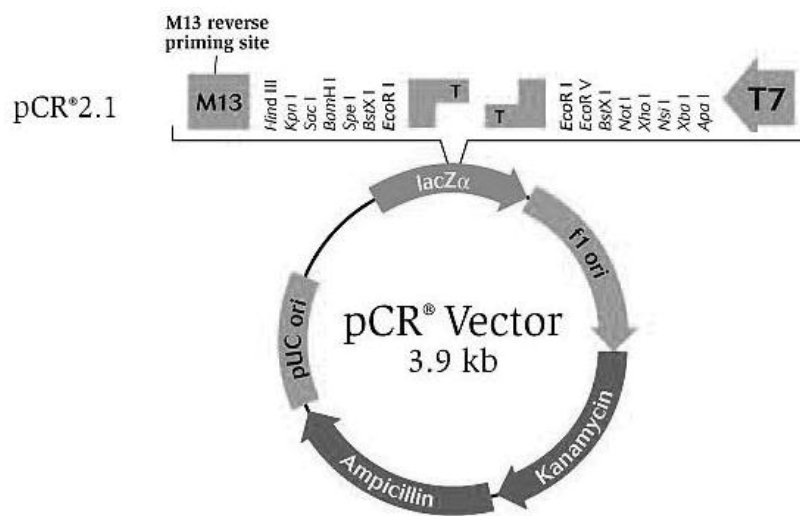


Figura 1: Mapa do vetor pCR 2.1, mostrando os s cios de restriç o (TA Cloning kit, Invitrogen).

Ap s quantificaç o do inserto no gel de agarose 1%, foi realizada a ligaç o a 16 C, durante 16 horas, de acordo com a seguinte re a o:

- Inseto (produto de PCR): 150 ng (1  l)
- Vetor pCR 2.1 TA: 50 ng (2  l)
- Tamp o T4 DNA ligase 10X: 1  l
-  gua DNase "free": 5  l
- T4 DNA ligase: 1  l

Em seguida, o produto da ligaç o foi transformado em bact rias *E. coli* TOP10 quimiocompetentes.

3.3.6 - Preparo de bactérias quimiocompetentes

Um inóculo de 5 ml contendo bactérias *E. coli*, cepa TOP10, foi cultivado em meio de cultura LB, overnight, a 37°C, velocidade média (140 rpm). No dia seguinte, foi feita uma diluição 1:100 do inóculo em LB, que cresceu em velocidade alta (200 rpm) até a OD₆₀₀ estar entre 0.4 e 0.6. Quando a OD₆₀₀ atingiu o valor adequado, a cultura foi separada em tubos de centrífuga gelados, resfriando-os no gelo por 15 minutos. A partir desse ponto todas as etapas foram realizadas com incubação constante das amostras no gelo quando não estavam sendo manipuladas. Em seguida, centrifugou-se os tubos a 2.500 rpm, durante 10 minutos, a 4°C.

Após centrifugação, o sobrenadante foi retirado com cuidado e ressuspenso gentilmente em TfbI, com 40 % do volume dessa solução em relação ao volume de cultura de bactéria. Os tubos foram incubados no gelo durante 30 minutos e após esse tempo, centrifugados a 2.500 rpm, por 10 minutos, a 4°C. Depois dessa segunda centrifugação, o sobrenadante foi retirado com cuidado, e ressuspenso gentilmente em TfbII, correspondente a 4% do volume de cultura de bactéria. Assim que colocado TfbII, as células foram imediatamente aliqüotadas (50 ou 100 µl por “ependorf”) e incubadas em banho de gelo seco/etanol (ou nitrogênio líquido) até serem transportadas para o freezer – 80°C. As alíquotas foram guardadas a essa temperatura até serem utilizadas.

3.3.7 - Transformação de bactérias quimiocompetentes

Para realizar a transformação, uma alíquota de células *Escherichia coli* TOP10 quimiocompetentes, preparadas como descrito no item 3.3.6, foi retirada do freezer – 80°C e mantida no gelo por no máximo 15 minutos. A esta alíquota foram adicionadas de 1 a 3 µl do produto de ligação, que foi homogeneizada com agitação leve e incubada em banho de gelo por 30 minutos. Após incubação a alíquota era submetida a um choque térmico por 90 segundos, a 42°C e rapidamente resfriada em banho de gelo por 2 minutos.

Logo em seguida foi adicionado meio de cultura LB e as bactérias foram incubadas a 37°C para se recuperarem do choque térmico, com agitação constante e rápida (200 rpm), de 45 minutos a uma hora. Depois desse período 50 e 100 µl de

bactérias foram plaqueadas em 20 ml de meio de cultura sólido LB ágar com ampicilina (100 µg/ml de ampicilina, IPTG 0,5 mM e X-gal 80 µg/ml).

Os meios de cultura sólidos contendo as bactérias foram incubados overnight, em estufa a 37°C, até que colônias pequenas e individualizadas aparecessem.

3.3.8 - Seleção de clones positivos

As colônias foram selecionadas através da coloração apresentada após incubação. O vetor pCR 2.1 apresenta dentro do sítio múltiplo de clonagem a sequência codificante para o peptídeo alfa da enzima β-galactosidase, enzima que, ao hidrolizar X-gal (um análogo da lactose), produz um precipitado de cor azul nas bactérias. Quando clonado, o inserto interrompe a sequência da enzima β-galactosidase. Sendo assim, os clones positivos são selecionados pela cor branca das colônias.

3.3.9 - Extração de DNA plasmidial (“mini-prep”)

A extração de DNA plasmidial das colônias positivas foi realizada de acordo com o protocolo adaptado de Sambrook *et al.*, 1989. As colônias foram cultivadas em 5 mL de meio LB líquido contendo 100 µg/ml de ampicilina, a 37°C, overnight. Após este período, 1 ml da cultura foi centrifugado a 12.000 rpm, por 6 minutos, à temperatura ambiente e o sobrenadante descartado. Essa operação foi repetida 2 vezes. Ao final, o pellet correspondente a 3 ml de cultura bacteriana foi dissolvido com 100 µl da solução de ressuspensão gelada. A mistura foi agitada até que se obtivesse dissolução completa do pellet e, em seguida, era incubada à temperatura ambiente por 5 minutos.

Após essa etapa, adicionou-se 200 µl da solução de lise. Os tubos foram misturados por inversão até que a amostra ficasse transparente, e incubados no gelo por 5 minutos.

Após esse período foram adicionados 150 µl da solução de neutralização, os tubos foram misturados por inversão durante 10 segundos e incubados no gelo por 5 minutos, seguido por uma centrifugação a 14000 rpm, durante 5 minutos, a temperatura ambiente. O sobrenadante foi transferido para um tubo novo e

precipitado com o dobro do volume de etanol absoluto. A nova mistura foi incubada por 30 minutos a -80°C ou overnight a -20°C , sucedido pela centrifugação de 14000 rpm, 15 minutos, 4°C . O sobrenadante foi descartado e o pellet lavado com 50 μl de etanol 70%. O sobrenadante foi novamente descartado, o pellet seco em speed vac e ressuspensionado em tampão TE ou água com RNase a uma concentração final de 20 $\mu\text{g}/\text{ml}$, sendo incubado por 40 minutos a 37°C .

Logo em seguida, o DNA foi extraído com o mesmo volume de fenol, centrifugado a 14000 rpm, a temperatura ambiente. A parte aquosa dessa mistura foi transferida a um novo tubo e acrescentado o mesmo volume de clorofórmio. A fração aquosa foi novamente retirada e adicionado o dobro de volume de etanol absoluto, precipitando-se o DNA overnight a -20°C ou 30 minutos em freezer -80°C . Essa solução foi em seguida centrifugada a 14000 rpm, 4°C , por 15 minutos, e o precipitado foi lavado duas vezes com etanol 70%, seco no speed vac e ressuspensionado em 30 μl de água livre de DNase.

De forma alternativa, os plasmídeos também eram obtidos utilizando o kit Wizard Plus mini-preps DNA purification System (Promega), seguindo protocolo estabelecido pelo fabricante.

Para verificar a quantidade e qualidade de DNA plasmidial obtido nesse processo as amostras foram submetidas a eletroforese em gel de agarose como descrito na seção 3.3.4. Os clones desejados foram escolhidos para as próximas etapas.

3.3.10 - Verificação da presença do inserto

3.3.10.1 - PCR com primers específicos

Para verificar a presença do inserto correspondente ao gene de CMS2MS2 foi realizada a mesma reação de PCR descrita na seção 3.3.3, com exceção de que o DNA utilizado na reação correspondia aos plasmídeos purificados por mini-prep.

3.3.10.2 - Digestão do vetor pCR 2.1 com enzimas de restrição

Para observar a liberação do inserto foi utilizada a enzima de restrição EcoRI, de acordo com o seguinte protocolo:

- DNA: 200 ng
- Tampão Eco RI 10 X : 1 μ l
- EcoRI : 1 μ l (1U/ μ l)
- Água DNase “free” q. s. p.: 10 μ l

A mistura foi incubada durante 3 horas, em banho-maria a 37°C e o produto da digestão foi analisado em gel de agarose 1% como descrito na seção 3.3.4.

A digestão dupla com as enzimas de restrição HindIII e XhoI também foi realizada alternativamente. O protocolo para digestão foi:

- DNA: 200 ng
- Tampão K 10 X: 1 μ l
- XhoI : 1 μ l (1 U/ μ l)
- HindIII: 1 μ l (1 U/ μ l)
- Água DNase “free” q.s.p.: 10 μ l

A mistura foi incubada por 3 horas, em banho-maria a 37°C. O produto da digestão foi analisado em gel de agarose 1% (seção 3.3.4).

3.3.11 - Sequenciamento dos clones positivos

Os clones que apresentaram a liberação do inserto do tamanho esperado e a amplificação de uma banda de aproximadamente 650 bp na reação de PCR foram sequenciados utilizando os primers “M13 forward” e “M13 reverse”, seguindo o protocolo para BigDye™ Terminator Ready Reaction Mix (derivado de Sanger *et al.*, 1977).

A reação de sequenciamento era montada da seguinte maneira:

- Big Dye: 0,33 μ l
- Tampão da reação: 0,77 μ l
- DNA: 200 ng
- Primer (M13 forward ou M13 reverse): 0,8 μ l (2 pmol)
- Água DNase "free" q. s. p.: 10 μ l

Essa mistura era submetida a PCR seguindo o programa de temperaturas:

- 1) 96°C, 2 minutos
- 2) 96°C, 30 segundos
- 3) Rampa de - 1 °C/s
- 4) 50°C, 15 segundos
- 5) Rampa de + 1 °C/s
- 6) 72°C, 4 minutos
- 7) Repetir de 2 a 6, 44 vezes
- 8) 4°C infinito

Após PCR, era adicionada às amostras 40 μ l de isopropanol 65%, esperando-se 30 minutos. Os tubos eram centrifugados 14000 rpm durante 25 minutos, o sobrenadante era retirado e 300 μ l de etanol 60% eram adicionados. Os tubos eram centrifugados novamente a 14000 rpm durante 5 minutos, o sobrenadante retirado e o pellet era seco ao ar livre.

Logo antes do sequenciamento, eram adicionados às amostras 20 μ l de formamida, que então eram incubadas a 94°C por 3 minutos. Em seguida, as reações foram aplicadas no sequenciador automático 3.100 Genetic Analyzer (Applied Biosystems).

Os sequenciamentos também eram feitos no sequenciador automático MegaBace 1000 DNA Analysis System (GE HealthCare Lifesciences), em que 200 ng de DNA eram utilizados como material inicial da amplificação por PCR, que seguia os seguintes parâmetros:

- 1) 95°C, 20 segundos
- 2) 50°C, 15 segundos
- 3) 60°C, 90 segundos
- 4) Repetir de 1 a 3, 29 vezes
- 5) 4°C infinito

No tubo de reação eram adicionados 4 µl de ET Kit (GE Healthcare Lifesciences) para 10 µl de volume final. Após a reação, foi acrescentado acetato de sódio 3M para precipitação do DNA, que foi ressuspenso em tampão do kit de sequenciamento. Em seguida as reações eram aplicadas no sequenciador.

As sequências obtidas nos dois tipos de sequenciamentos eram visualizadas no programa SeqTools, versão 8.4 (S. W. Rasmussen) ou Chromas (Technelysium), onde eram analisadas quanto à sua qualidade.

3.3.12 - Análises *in silico*

As sequências de DNA obtidas no sequenciamento do vetor pCR 2.1 contendo o inserto foram analisadas *in silico* utilizando ferramentas como “blastn” e “blastx” disponíveis no servidor do banco de dados NCBI (“National Center for Biotechnology Information”, disponível em: <http://www.ncbi.nlm.nih.gov>).

A tradução *in silico* da sequência nucleotídica também foi feita utilizando ferramentas presentes no servidor Expasy (“Expert Protein Analysis System”, disponível em: <http://www.expasy.ch>). O alinhamento de proteínas foi realizado com o software Clustalw EBI 2.0, também do servidor Expasy (disponível em: <http://www.ebi.ac.uk/Tools/clustalw2/index.html>). A sequência aminoacídica obtida foi alinhada com cisteíno-proteases de Caricaceae, em especial com a sequência de CMS2MS2 disponibilizada por Gomes e colaboradores (2007). A análise dos alinhamentos foi feita utilizando o diagrama de Venn, no qual o programa Clustalw EBI se baseia para definir que tipo de conservação de aminoácidos ocorre entre as sequências alinhadas.

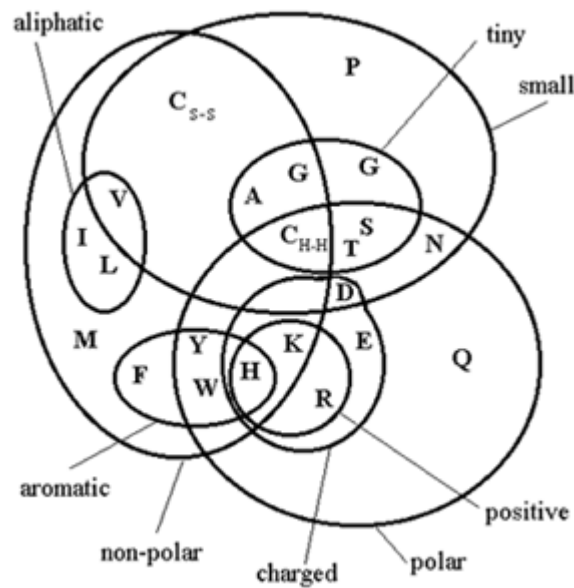


Figura 2: Diagrama de Venn, utilizado pelo programa Clustalw EBI para verificar a natureza das substituições de aminoácidos.

Outros parâmetros seguidos pelo programa Clustalw EBI estão disponíveis em: <http://www.ebi.ac.uk/Tools/clustalw2/help.html#results>.

A verificação de sítios de glicosilação nas sequências foi feita utilizando o programa NETNGLYC, também disponível no servidor ExPASy (disponível em: <http://www.cbs.dtu.dk/services/NetNGlyc/>).

3.3.13 - Purificação do inserto e obtenção do vetor pET28a(+)-TEV

Uma vez confirmada a presença do gene no vetor pCR 2.1, uma nova digestão com maiores quantidades de DNA foi realizada (segundo descrito na seção 3.3.10.2). Essa digestão foi realizada para obter quantidades suficientes de inserto puro para a ligação.

O produto da digestão foi aplicado em gel de agarose 1% do tipo “low-melting”. O inserto foi excisado e purificado, utilizando o kit GFX PCR DNA and Gel Band Purification Kit (GE HealthCare Lifesciences), segundo protocolo do fabricante. Após a extração, a quantidade de DNA foi determinada por espectrofotometria a A_{280}/A_{260} e por análise em gel de agarose 1%.

O plasmídeo pET28a(+)-TEV, por sua vez, foi obtido por “mini-prep” (seção 3.3.9). O vetor pET28a(+)-TEV é uma modificação do vetor pET28a(+), tendo seu

sítio de clivagem por trombina trocado por um sítio de clivagem pela proteína TEV (ref). A figura 3 mostra um esquema desse vetor.

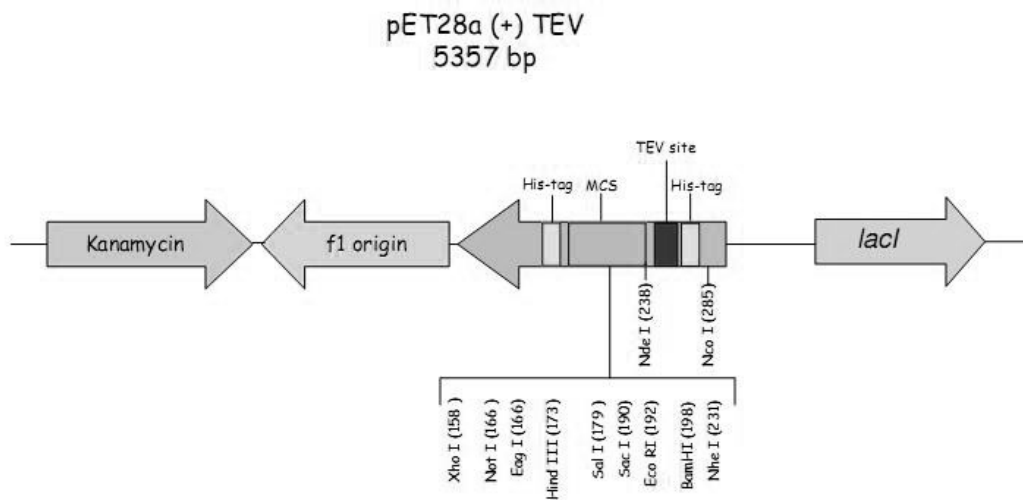


Figura 3: Mapa do vetor pET28a(+)-TEV, mostrando os sítios de restrição.

O vetor foi digerido com as enzimas Sall e HindIII, escolhidas para a sub-clonagem. Sall gera um produto com pontas coesivas ao produto gerado por XhoI. A digestão realizada por Sall ocorreu de forma separada à de HindIII, devido à incompatibilidade de funcionamento dessas enzimas em um mesmo tampão.

A digestão do vetor pET28a(+)-TEV pela enzima de restrição Sall, overnight, seguiu o seguinte protocolo:

- DNA : 2 µg (aproximadamente 50 µl)
- Tampão M 10 X: 10 µl
- NaCl 1 M: 10 µl
- BSA: 10 µl
- Sal I: 7 µl (1U/µl)
- Água DNase "free" q.s.p.: 100 µl

Após a digestão com Sall, o produto da digestão foi purificado em mini-coluna S-200 previamente montada, aplicando-se o DNA sobre a resina e centrifugando a 14000 rpm por 2 minutos, a temperatura ambiente.

O produto da purificação era então submetido a uma nova digestão, dessa vez com HindIII, durante 3 horas, seguindo o protocolo:

- DNA : 1 µg (aproximadamente 60 µl)
- Tampão K 10 X: 10 µl
- Sal I: 10 µl (1U/µl)
- Água DNase “free” q.s.p.: 100 µl

O produto dessa digestão foi aplicado em gel de agarose 1 %, excisado e purificado como em 3.3.13 .O vetor digerido por Sall e HindIII foi então quantificado em gel de agarose 1% e por espectrofotometria a A_{280}/A_{260} .

3.3.14 - Ligação do inserto ao vetor pEt28a(+)-TEV e transformação em bactérias *E. coli* BL21-DE3

A ligação do inserto ao vetor pET28a(+)-TEV foi realizada como em 3.3.5, usando T4 DNA ligase (Invitrogen), na proporção de 3 de inserto: 1 de vetor, a 16 °C, overnight.

A ligação foi transformada em bactérias *Escherichia coli* BL21-DE3 quimiocompetentes, como descrito na etapa 3.3.7. O cultivo dessas bactérias em meio sólido foi realizado como em 3.3.7, com exceção que não foram acrescentados IPTG e X-gal ao meio de cultura, pelo fato de que pET28a(+)-TEV não possui a enzima da β-galactosidase em seu sítio de múltipla clonagem. A análise por PCR em busca de colônias positivas foi realizada como em 3.3.10.1.

Alguns clones foram escolhidos para serem sequenciados utilizando os primers correspondentes ao vetor pET28a(+)-TEV: primers T7 “promoter” e T7 “terminator”. Esses primers, assim como os primers M13, se anelam próximo ao inserto ligado, fornecendo informações sobre a fase de leitura na qual se encontra a sequência e sua constituição, como presença de “stop” codons prematuros.

O sequenciamento dos clones foi realizado como na seção 3.3.11 e a análise das sequências obtidas como em 3.3.12.

3.3.15 - Expressão de CMS2MS2 recombinante

As colônias positivas para CMS2MS2 foram cultivadas em meio LB overnight, e em seguida uma alíquota da cultura foi cultivada em meio LB na diluição de 1:100, a 37°C, em velocidade rápida (200 rpm), até que atingisse a OD₆₀₀ de 0.6 a 0.8, quando 0,5 mM de IPTG foi acrescentado à cultura. Alíquotas de 1 ml foram coletadas de hora em hora, até que se completou 4 horas de expressão. As alíquotas foram processadas para que houvesse lise das bactérias e liberação do conteúdo citoplasmático.

Nesse procedimento, a cultura de bactéria era centrifugada a 4000 rpm, durante 20 minutos, a 4°C. O sobrenadante dessa primeira centrifugação era descartado. Em seguida, o pellet foi dissolvido em PBS 1X gelado contendo 0,25 % de lisozima e essa mistura foi submetida a três choques térmicos sucessivos em banho-maria 37°C/ nitrogênio líquido -70°C. Esse procedimento teve por objetivo lisar as bactérias para liberação do conteúdo citoplasmático. Depois, os pellets dissolvidos foram sonicados, na amplitude 30, com ciclos de 30 segundos sonicação/30 segundos intervalo, por três vezes, evitando a formação de espuma. Durante toda a lise as amostras foram mantidas em gelo quando não manipuladas. Após duas etapas de lise, os tubos foram centrifugados a 10000 rpm, a 4 °C, durante 30 minutos. O sobrenadante dessa centrifugação correspondia à parte solúvel da expressão, enquanto que o pellet à parte insolúvel. A parte solúvel, portanto, foi separada para análise em gel de poliacrilamida.

A parte insolúvel foi processada para que houvesse a dissolução de proteínas insolúveis, presentes em corpos de inclusão da bactéria, ou para dissolver proteínas agregadas. Ao pellet foi adicionada uréia a 8M e dissolvido até que ficasse homogêneo. Essa solução foi centrifugada a 10000 rpm, 4°C, durante 20 minutos. O sobrenadante dessa última centrifugação corresponde à fração insolúvel da expressão. O pellet, que possuía materiais altamente insolúveis da bactéria como a parede celular, foi descartado.

As frações solúvel e insolúvel, portanto, foram aplicadas em gel de poliacrilamida com o objetivo de verificar a expressão da proteína recombinante.

3.3.16 - Eletroforese em Gel de Poliacrilamida Desnaturante:

Utilizou-se o sistema de eletroforese em gel de poliacrilamida desnaturante com SDS [“Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis” (SDS-PAGE)] descrito por Laemmli (1970). Esse sistema contém um gel separador (item 3.1.2), que foi adicionado a uma placa de vidro com 8,0 x 10,0 x 0,3 cm de dimensões, com espaçadores e vedantes de borracha. Após a polimerização desse gel, preparou-se o gel concentrador (item 3.1.2) adicionando-o sobre o gel separador e inserindo um pente com 10 dentes sobre esse último gel. Depois de polimerizado, o pente e os vedantes foram cuidadosamente retirados, sendo a montagem colocada na cuba de eletroforese contendo o tampão de corrida glicina 1 X (item 3.1.2).

Foi adicionado um volume do tampão de amostra (item 3.1.2), correspondente a 1/6 do volume total. Nas amostras contendo proteases também foi adicionado 1 µl de MMS 50 mM anterior à fervura por 5 min para a desnaturação das mesmas. A aplicação foi feita nas canaletas localizadas no pólo negativo do gel. A eletroforese foi desenvolvida a 80 V de tensão, 20 mA de corrente e 2 W de potência, durante aproximadamente 2 h a 7°C ou temperatura ambiente.

Terminada a corrida eletroforética, transferiu-se o gel para a solução fixadora (item 3.1.2) na qual permaneceu por 15 min. Depois foi adicionada a solução de azul-brilhante de coomassie para coloração de géis (3.1.2) a qual permanecia por 30 minutos ou até o aparecimento das bandas. O fundo do gel (“background”) foi descorado com a solução descorante preparada como descrito no item 3.1.2.

A foto digital do gel foi tirada em câmera Kodak DC-40, sobre um transiluminador para a análise dos resultados. A intensidade das bandas foi calculada por densitometria usando o programa computacional Kodak versão 1D-3.5.

A massa molecular foi estimada através da análise do gel desnaturante corrido com padrões de peso molecular com massas conhecidas: lisozima, 14 kDa, a-quimiotripsinogênio, 25 kDa e ovalbumina 45.000 kDa. O cálculo do tamanho da banda de interesse foi realizado utilizando o programa computacional Kodak versão 1D-3.5.

3.3.17 - Análise de parâmetros bioquímicos

Confirmada o tamanho da banda de expressão com o que era esperado, e utilizando a tradução *in silico* de aminoácidos, alguns parâmetros bioquímicos teóricos da proteína foram calculados. Foi utilizado o programa ScanSite, disponível no servidor Expasy (http://scansite.mit.edu/calc_mw_pi.html) para análise da massa molecular (MW) e do ponto isoelétrico (pI) da proteína recombinante.

3.3.18 - Purificação por cromatografia de afinidade ao níquel

O tempo de expressão de 4 horas foi escolhido e uma nova expressão da proteína recombinante foi realizada, dessa vez para coletar material para purificação.

A fração escolhida foi a insolúvel, já que não foi encontrada proteína recombinante na fração solúvel do lisado bacteriano. A fração insolúvel, portanto, foi submetida à cromatografia de afinidade ao níquel na coluna Ni Sepharose™ High Performance (GE HealthCare Lifesciences).

Previamente, a coluna, de dimensões 13 cm de altura por 0,5 cm de diâmetro, foi equilibrada com tampão de ligação (preparado conforme seção 3.1.2) com ajuda de uma bomba peristáltica ligada a um medidor de absorvância a 280 nm. A fração insolúvel também foi dissolvida neste tampão, que continha 20 mM de imidazol. Nessa concentração, o imidazol força a eluição de proteínas do lisado que possuam fraca afinidade ao níquel, mas ela é pouco o suficiente e a proteína recombinante se mantém ligada. Essa mistura foi aplicada à coluna e, logo em seguida, lavou-se a coluna com tampão de ligação afim de eluir as proteínas que não possuíam afinidade à coluna. Essa lavagem durou até que picos a 280 nm não fossem mais observados e se atingisse a linha basal de absorção observada no início da eluição.

Após essa etapa, aplicou-se tampão de eluição à coluna, contendo 250 mM de imidazol. A saída de proteínas foi observada pelo contador de absorvância, embora o imidazol tenha mascarado a detecção da saída da proteína recombinante, pois essa molécula também possui absorvância a 280 nm.

Várias frações foram coletadas, separando principalmente as coletas depois da aplicação do tampão de eluição. As amostras coletadas foram analisadas em gel de poliacrilamida 12% para presença da proteína fusionada à cauda de histidina.

3.3.19 - Ensaios de atividade mitogênica com CMS2MS2 recombinante

3.3.19.1 - Esterilização das amostras para ensaios biológicos

Antes dos experimentos as amostras foram esterilizadas por filtração em membranas de nitrato de celulose, com 0,22 µm de diâmetro, a fim de evitar a contaminação.

3.3.19.2 - Cultivo das linhagens celulares:

Células L929 foram cultivadas em RPMI 1640 (preparado como descrito no item 3.1.2) contendo 5% (v/v) de soro fetal bovino (FBS), enquanto que osteoblastos de cultura primária foram cultivados com este meio suplementado com FBS 10% e gentamicina 100 mg/ml. As linhagens foram mantidas em frascos de cultivo na estufa de cultura a 37°C, com atmosfera úmida e 5% (v/v) de CO₂.

Ao atingirem aproximadamente 50% de confluência, foram realizados sub-cultivos para os ensaios biológicos. Células L929 e osteoblastos eram inicialmente lavadas com PBS/EDTA pH 7,4 (preparado como descrito no item 3.1.2) e em seguida adicionada tripsina 0,2% (v/v) diluída em RPMI. Quando as células apresentavam-se desprendidas da superfície do frasco de cultivo, foi adicionado meio RPMI com FBS 5%. A contagem das células obtidas após tripsinização era feita utilizando câmara de Neubauer.

3.3.19.3 - Medida da atividade mitogênica das amostras através da incorporação de [H³-metil]-timidina no DNA

Células da linhagem L929 e osteoblastos de cultura primária foram semeadas em placas de poliestireno de 24 cavidades (1,5 x 10⁴ células/cavidade), suspensas em 1,0 ml de meio RPMI com FBS 5%, sendo incubadas posteriormente em estufa de cultura até atingirem 50% de confluência. As células, então, foram carenciadas da presença de fatores de proliferação através da incubação destas com meio RPMI com FBS 0,2%. Tendo decorrido 24 h de carenciamento, foram adicionadas as amostras esterilizadas diluídas em meio de cultura RPMI. Além do FBS 0,2 %, como

controle em que há estímulo mínimo por fatores de crescimento, foi mantido um controle positivo de proliferação com FBS 6%, ambos diluídos em RPMI.

Após 24 h de exposição às amostras, foi adicionado 1,0 μCi /cavidade de [H^3 -metil]-timidina, deixando as placas em condições de cultivo por 18 h. Em seguida, foi aspirado o meio de cultura e as células foram lavadas duas vezes com 1,0 ml de meio RPMI. As células, então, foram fixadas com duas lavagens utilizando 1ml da solução de TCA 5% (v/v). Para desidratá-las utilizou-se 1 ml de etanol 95% (v/v), deixando as placas secarem ao ar livre. Depois as células foram dissolvidas com 250 μl de hidróxido de sódio 0,5 M, neutralizando em seguida com a adição de 375 μl de ácido clorídrico 1,0 M. Retirou-se o líquido contendo as células dissolvidas, que foram colocadas em tubos de polipropileno, adicionando a cada um 3,5 ml de aquassol (preparado como descrito no item 3.2.1). Depois foi feita a contagem automática das cintilações por minuto (CPM), sendo o resultado expresso convertendo-se os dados das CPM obtidas para “índice de proliferação”. Para isso, foi atribuído à média das CPM das cavidades tratadas com FBS 0,2% um valor arbitrário de 100%. A análise dos dados foi realizada no programa computacional Sigma Plot versão 5.0 (Sytat Software Inc).

Na avaliação da atividade mitogênica em osteoblastos, as células foram semeadas em meio de cultura RPMI suplementado com FBS 10% com gentamicina 100 mg/ml. Os osteoblastos foram carenciados em FBS 1%, no período de 16 h. O controle negativo também era FBS 1%. As demais etapas foram idênticas às descritas para L929.

4. Resultados

4.1 - Amplificação de CMS2MS2 e clonagem em vetor pCR 2.1

A clonagem do cDNA da cisteíno-protease de *Carica candamarcensis* correspondente à enzima CMS2MS2, cuja sequência havia sido descrita previamente por Gomes *et al.* (2007), foi iniciada com o isolamento do RNA de folhas da planta.

A extração de RNA foi realizada como descrito no item 3.2.1. A análise do RNA total em gel de agarose/formaldeído (seção 3.2.1) é mostrado na figura 4.

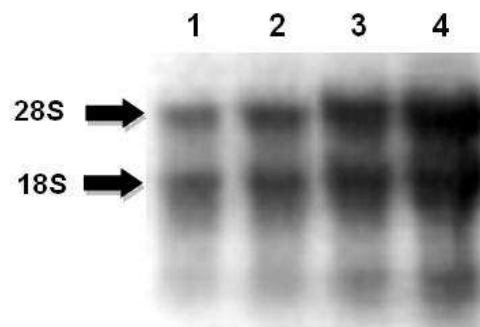


Figura 4: Eletroforese em gel de agarose/formaldeído de RNA total extraído de folhas de *C. candamarcensis*. De 1 a 4, concentrações crescentes (25, 50, 100, 150 µg) de RNA total. As setas correspondem às subunidades ribossomais 28S e 18S.

A síntese de DNA complementar foi realizada de acordo com o descrito no item 3.2.2. O RNA total, previamente extraído e armazenado seco a -70°C , foi dissolvido para a concentração final de $25\ \mu\text{g}/\mu\text{l}$ e utilizado como molde para a síntese do cDNA através da reação de RT-PCR (seção 3.2.2). O DNA complementar obtido foi utilizado para a amplificação do fragmento de interesse através de PCR convencional.

Como base na sequência de aminoácidos de CMS2MS2, desvendada por Gomes *et al.* (2007), dois primers que se anelam à região correspondente às porções N e C-terminal da proteína foram desenhados. A execução da PCR utilizando esse par de primers (seção 3.3.3) gerou dois fragmentos principais, apresentando tamanhos de aproximadamente 500 e 650 bp quando o produto da amplificação por PCR foi analisado em gel de agarose 1%, corado com solução de brometo de etídeo (figura 5).

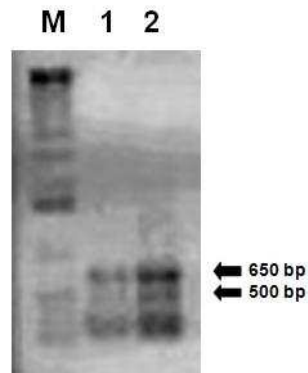


Figura 5: Eletroforese de fragmento amplificado a partir de cDNA utilizando os primers CMS2MS2. Gel de agarose 1% corado com brometo de etídeo. M: padrão de peso molecular 1 kb DNA Plus ladder (Invitrogen). 1 e 2: produtos de amplificação de cDNAs diferentes.

A banda mais intensa na canaleta 2 da figura 5 corresponde à banda de 650 bp. Abaixo da banda de 500 bp existem outras bandas, de menor tamanho molecular, que correspondem aos primers. Dois estoques de cDNAs diferentes foram utilizados nas reações de PCR que deram origem aos produtos presentes nas canaletas 1 e 2, que possuem perfil semelhante.

Tomando como referência o tamanho da proteína CMS2MS2 que possui 214 aminoácidos, estimava-se que o fragmento de aproximadamente 650 bp fosse o da proteína CMS2MS2. Desse modo, o fragmento de 650 bp foi excisado do gel de agarose, purificado como descrito em 3.3.13 e quantificado novamente por análise em gel de agarose 1% com o marcador λ DNA Hind III. O produto da purificação foi utilizado para a ligação com o vetor pCR 2.1.

Bactérias quimiocompetentes *E. coli* TOP10 foram transformadas com o produto da ligação e plaqueadas em meio LB ágar contendo 100 μ g/ml de ampicilina, IPTG 0,5 mM e X-gal 80 μ g/ml. As colônias positivas (escolhidas pela coloração branca) foram submetidas à extração de DNA plasmidial (seção 3.3.9). Os DNAs plasmidiais foram analisados em gel de agarose 1% (figura 6).

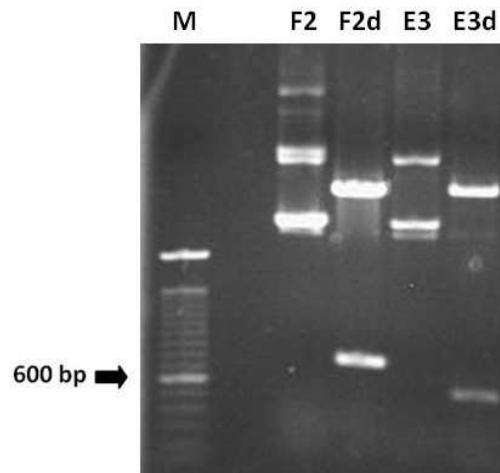


Figura 6: Eletroforese de DNA plasmidial pCR 2.1 dos possíveis clones de CMS2MS2, F2 e E3. Gel de agarose 1% corado com brometo de etídeo. M: padrão de massa molecular 100 bp DNA ladder (Invitrogen). F2: mini-prep do plasmídeo F2 (aproximadamente 300 ng de DNA). F2d: plasmídeo F2 digerido com Eco RI (300 ng). E3: mini-prep do plasmídeo E3 (100 ng). E3d: digestão do plasmídeo E3 com HindIII e XhoI (100 ng).

Este gel demonstra que, após digestão por enzima de restrição, o clone F2 liberou um fragmento entre 600 e 700 bp. Esse fragmento corresponde à digestão com a enzima EcoRI, que possui sítios de restrição flanqueando a região de inserção do gene no plasmídeo pCR 2.1. Entretanto, a digestão com enzimas de restrição HindIII e XhoI do clone E3 evidenciou uma banda de 400-500 bp, tamanho menor que o aguardado. Isso sugere que o inserto possuía sítio de restrição para uma dessas enzimas. Essa possibilidade foi confirmada após sequenciamento do clone E3, que mostrou a existência de um sítio de restrição para Hind III (figura 8).

4.2 - PCR, sequenciamento dos clones e análise *in silico*

Após confirmada a presença do inserto do tamanho esperado, os plasmídeos recuperados por mini-prep foram submetidos a sequenciamento de DNA utilizando os primers universais M13 “forward” e “reverse”. Vários clones apresentaram sequenciamento positivo para cisteíno-proteases semelhantes a CMS2MS2. Entretanto, somente E3 e F2 serão apresentados a partir de agora nesse trabalho, devido à qualidade superior dos sequenciamentos desses clones.

Os sequenciamentos foram feitos utilizando ora o primer M13 “forward”, ora o primer M13 “reverse”. Os sequenciamentos foram avaliados quanto à sua qualidade

no programa SeqTools ou Chromas (seção 3.3.11) e no mínimo quatro deles foram analisados a fim de confirmar a autenticidade das sequências fornecidas pelo sequenciador automático. O resultado das sequências consenso dos clones denominados E3 e F2 são mostrados nas figuras 7 e 8.

```
5'...GTTTCCCCAGTCAACGACGTTGTAAACGGACGGCCAGTGAATTGTTAATACGACTCACTATAGGGCGAAT
TGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGGCTTTTACCCAGA
GAGTATTGATTGGAGACAGAAAGGAGCTGTCACTCCTGTA AAAAGATCAAATCCTTGCGGTAGTTGTTGGGC
ATTTTCAACTGTTGCAACTGTAGAGGGAATAAACAAGATTGTA ACTGGAAAGTTAATATCATTATCAGAACAAG
AATTGCTAGATTGTGACAGACGTAGCCACGGGTGTAATGGAGGTTACCAAACGACATCACTTCAATATGTGG
TTGACAATGGTGTTCACACGGAATATGAGTATCCTTACGAGAAGAAGCAAGGCAATTGTCGAGCCAAAGATA
AGAAAGGTCTGAAAGTGCAAATTA CTGGGTATAAACGTGTGCCACCAAATGATGAAATATCTCTATAAAAGT
CATTGCAAATCAACCTGTGAGCGTCTCATTGAGTCCAAAGACAGAAGTTTCCATTCTACAGAGGGGGAAATT
TATAAAGGGCCATGCGGAACTAGATTAGATCATGCAGTAACAGCAATCGGGTACGGAAAAGATTACATACTC
ATAAAGAATTCATGGGGTCCAAACTGGGGCGAGAAGGGATATATAAGAATTA AAAGAGCATCTGGAAAATCC
GAGGGAATATGTGGAGTTTATAAGAGCTCATACTTCCTATTA AAGGTTACCAAAAGCCGAATTCCAGCACA
CTGGCGGCCGTTACTAGTGGATCCCGAGGCTCGGGTACCTAAAGGA...3'
```

Figura 7: Sequência de nucleotídeos do clone F2, obtida em sequenciador automático 3.100 Genetic Analyzer (Applied Biosystems), utilizando os primers M13 “forward” ou “reverse”. Em negrito e sublinhado, a sequência dos primers que foram utilizados na amplificação do gene putativo da proteína CMS2MS2.

```
5'...CAGAACATAAAGGGGCGAATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCT
GCAGAATTCGGCTTTTATCCTGAAAGTATCGACTGGAGAAAGAAAGGAGCTGTCACTCCTGTAAAAAATCAAG
GTTCTTGCGGTAGTTGTTGGGCATTTTCAACTATTGCAACAGTGGAGGGAATAAACAAGATTGTACATGGGAA
TTTAACATCATTATCGGAGCAAGAATTGGTAGATTGTGACAGACGTAGCCACGGGTGTA AAGGAGGTTACCA
AACGACCTCACTTAAATATGTGGTTGACCATGGTGTTCACACGGA AAAAGGAGTATCCATACGAACAGAAGCA
ATGCAAGTGTGAGCTAAAGATAAGAAACCTCCGATAGTAAAAATTAGTGGGTATAAACGTGTGCCATCAAAT
GATGAAATATCTCTATAAAAGCAATTGCAAACAACCTGTGAGCGTCTTGTGTAATCTACAGGAAAAGCTT
TCCAATTCTATAAAAGGGGATTTTTGGAGGGCCATGCGGAACTAAAGTAGATCATGCAGTAACAGCTGTGCG
GTTACGGAAAAGATTATATACTCATAAAGAATTCATGGGGTCCAAAATGGGGCGATAATGGATACATAAAAAA
TTAAAGAGCAAAGTGGGACACTGCGAGGGAATACGTGGAATTTATAAAGAGCTCA TATTTCCCAATTA AAGG
CTATCAAAAGCCGAAGTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCG
TATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACAACATACGAGCCGGAAG
CATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTTCACATTAATTGCGTTGCGCTCACTGCCCGG
CTTTCCAGTCAGCAACCTGTCTGCAGCTGCATTAATGAATCGGCCACGCGCGGGGAAGAGGCGGTTAGCGT
ATTGGGCGCTCTTCGA...3'
```

Figura 8: Sequência de nucleotídeos do clone E3, obtida em sequenciador automático 3.100 Genetic Analyzer (Applied Biosystems), utilizando os primers M13 “forward” ou M13 “reverse”. Em negrito e sublinhado, a sequência dos primers que foram utilizados na amplificação do gene putativo da proteína CMS2MS2. Em cinza são mostrados os sítios internos de HindIII (AAGCTT) e XhoI (CTCGAG) e, sublinhado, um “stop” códon.

A identidade entre as sequências de DNA dos clones F2 e E3 é de 86 %.

As sequências de nucleotídeos de E3 e F2 foram traduzidas *in silico* usando o programa “Translate Tool”, do servidor Expsy (seção 3.3.12). As sequências de aminoácidos são apresentadas a seguir.

```
(...)YPESIDWRQKGAVTPVKDQNPCGSCWAFSTVATVEGINKIVTGKLI
SLSEQELLDCDRRSHGCNGGYQTTSLQYVVDNGVHTEYEYEPYEKKQG
NCR AKDKKGLKVQITGYKRVPPNDEISLIKVIANQPVSVLIESKDRSFH
FYRGGIYKGPCGTRLDHAVTAIGYGKDYILIKNSWGPNWGEKGYIRIK
RASGKSEGICGVYKSSYFPIKGYQ(...)
```

Figura 9: Sequência da proteína hipotética da região do inserto do clone F2 correspondente à região codificante de CMS2MS2. (...) representa a região do vetor pCR2.1.

```
(...)YPESIDWRKKGAVTPVKNQGSCGSCWAFSTIATVEGINKIVHGNLT
SLSEQELVDCDRRSHGCKGGYQTTSLKYVVDHGVHTEKEYPYEQKQC
KCR AKDKKPPIVKISGYKRVPSNDEISLIKAIKQPVSVLVESTGKAFAQ
FYKKGIFGGPCGTKVDHAVTAVGYGKDYILIKNSWGPKWGDNGYIKN
StopKSKVGHCEGIRGIYKSSYFPIKGYQ(...)
```

Figura 10: Sequência da proteína hipotética da região do inserto do clone E3 correspondente à região codificante de CMS2MS2. (...) representa a região do vetor pCR2.1.

De acordo com a previsão do programa NETNGLYC, disponível no servidor Expsy (seção 3.3.12), o clone E3 possui dois sítios putativos de glicosilação em sua sequência, um em ${}_{44}\text{NLT}_{46}$ e outro em ${}_{189}\text{NST}_{191}$. Também segundo o programa, somente o primeiro sítio tem potencial para ser glicosilado.

Em seguida, as sequências de nucleotídeos de E3 e F2 foram submetidas à análise de homologia com sequências depositadas em bancos de dados, utilizando o algoritmo “blastn”, do programa Blast do NCBI (seção 3.3.12). Nas tabelas de 1 e 2 são mostrados os cinco maiores “scores” de identidade de sequências de DNA com a sequência dos clones F2 e E3.

NÚMERO DE ACESSO	DESCRIÇÃO DA SEQUÊNCIA	IDENTIDADE	Valor "e"
AY035986.1	<i>Carica candamarcensis</i> cysteine proteinase CC23 gene	94 %	1e-94
DQ836126.1	<i>Vasconcellea x heilbornii</i> VXH-C mRNA, partial cds	93 %	0
DQ836124.1	<i>Vasconcellea stipulata</i> VS-B mRNA, partial cds	91 %	0
DQ836124.1	<i>Vasconcellea stipulata</i> VS-A mRNA, partial cds	78 %	3e-81
DQ836122.1	<i>Vasconcellea x heilbornii</i> VXH-A mRNA, partial cds	78 %	8e-82

Tabela 1: Descrição dos cinco maiores "scores" de identidade da sequência de nucleotídeos F2 em relação à sequências depositadas em banco de dados, utilizando o algoritmo "blastn" do servidor NCBI. A identidade entre as seqüências foi priorizada em relação ao valor "e" durante a classificação decrescente apresentada na tabela.

NÚMERO DE ACESSO	DESCRIÇÃO DA SEQUÊNCIA	IDENTIDADE	Valor "e"
DQ836124.1	<i>Vasconcellea stipulata</i> VS-B mRNA, partial cds	90 %	0
DQ836126.1	<i>Vasconcellea x heilbornii</i> VXH-C mRNA, partial cds	87 %	4e-140
AJ131997.1	<i>Carica papaya</i> mRNA for chymopapain isoform IV, partial	82 %	2e-127
AJ131995.1	<i>Carica papaya</i> mRNA for chymopapain isoform II	81 %	2e-127
AJ131996.1	<i>Carica papaya</i> mRNA for chymopapain isoform III	81 %	5e-124

Tabela 2: Descrição dos cinco maiores "scores" de identidade da sequência de nucleotídeos E3 em relação à sequências depositadas em banco de dados, utilizando o algoritmo "blastn" do servidor NCBI. A identidade entre as seqüências foi priorizada em relação ao valor "e" durante a classificação decrescente apresentada na tabela.

A comparação da sequência de nucleotídeos de F2 e E3 com as cisteíno-proteases presentes no banco de dados do NCBI demonstrou alta identidade com as sequências das proteases VS-B, de *V. stipulata* e VXH-C, de *V. x heilbornii*. A sequência de DNA do clone F2 possui mais semelhança com as sequências gênicas de proteases da família *Vasconcellea* (CC-23, VS-A e VXH-A) enquanto E3 possui alta identidade com as isoformas de quimopapaína II, III e IV de *C. papaya*.

As sequências de nucleotídeos também foram analisadas na ferramenta "blastx", do servidor NCBI (seção 3.3.12). Nessa ferramenta, as sequências de DNA

são traduzidas em aminoácidos e é feito o alinhamento com sequências de proteínas depositadas no banco de dados do NCBI. Nas tabelas de 3 e 4 são mostrados os cinco maiores “scores” de identidade de sequências de aminoácidos com a sequência dos clones F2 e E3.

IDENTIFICAÇÃO	DESCRIÇÃO DA SEQUÊNCIA	IDENTIDADE	Valor “e”
gb ABI30274.1	VS-B <i>Vasconcellea stipulata</i>	86 %	6e-99
sp P84347.1 MEX2_JACME	Full Chymomexicain	84 %	1e-100
pdb 2BDZ A	Chain A, Mexicain from <i>Jacaratia mexicana</i>	83 %	1e-103
sp P84346.1 MEX1_JACME	Full Mexicain	82 %	5e-102
pir S46476	cysteine proteinase (EC 3.4.22.-) III	78 %	3e-96

Tabela 3: Descrição dos cinco maiores “scores” de identidade da sequência de aminoácidos de F2 com sequências depositadas em banco de dados, utilizando o programa “blastx” do NCBI. A identidade entre as seqüências foi priorizada em relação ao valor “e” durante a classificação decrescente apresentada na tabela.

IDENTIFICAÇÃO	DESCRIÇÃO DA SEQUÊNCIA	IDENTIDADE	Valor “e”
pir S46476	cysteine proteinase (EC 3.4.22.-) III	92 %	5e-111
gb ABI30274.1	VS-B <i>Vasconcellea stipulata</i>	83 %	1e-91
sp P84347.1 MEX2_JACME	Full Chymomexicain	74 %	2e-87
pdb 2BDZ A	Chain A, Mexicain from <i>Jacaratia mexicana</i>	73 %	1e-90
sp P84346.1 MEX1_JACME	Full Mexicain	72 %	8e-89

Tabela 4: Descrição dos cinco maiores “scores” de identidade da sequência de aminoácidos de E3 com sequências depositadas em banco de dados, utilizando o programa “blastx” do NCBI. A identidade entre as seqüências foi priorizada em relação ao valor “e” durante a classificação decrescente apresentada na tabela.

As sequências de aminoácidos de F2 e E3 são muito parecidas com as de mexicana e quimomexicana, cisteíno-proteases da espécie *Jacaratia mexicana*, que também pertence à família Caricaceae. A sequência traduzida de E3 possui alta identidade com CC-III, também de *C. candamarcensis*, o que não ocorre com a

sequência deduzida de F2, que codifica uma proteína muito semelhante a VS-B, de *V. stipulata*.

A sequência de aminoácidos traduzidas dos clone E3 e F2 foram alinhadas no programa ClustalW (figura 11).

```

F2      YPESIDWRQKGAVTPVKDQNPCGSCWAFSTVATVEGINKIVTGKLISLSEQELLDCDRRS 60
E3      YPESIDWRKKGAVTPVKNQGS CGSCWAFSTIATVEGINKIVHGNTLSLSEQELVDCDRRS 60
*****:*****:*. . *****:***** *: * *****:*****

F2      HGCNGGYQTTSLQYVVDNGVHTEYEYYPYEKKQGNCRAKDKKGLKVQITGYKRVPPNDEIS 120
E3      HGCKGGYQTTSLK YVVDHG VHTEKEYPYEQKQCKCRAKDKKPPIVKISGYKRVPSNDEIS 120
***:*****:****:***** *****:*** :***** *: **:*****.*****

F2      LIKVIANQPVS VLIESKDRSFHFYRGGIYKGPCGTRLDHAVTAIGYGKDYILIKNSWGPN 180
E3      LIKAI AKQPVS VLVESTGKAFQFYKKGIFGGPCGTRKVDHAVTAVGYGKDYILIKNSWGPK 180
***. **:*****:***. . :*:** :** :*****:*****:*****:*****:*****:

F2      WGEKGYIRI-KRASGKSEGICGVYKSSYFPIKGYQ 214
E3      WGDNGYIKN-KSKVGHCEGIRGLYKSSYFPIKGYQ 214
***:***: * *: . *** *:*****:*****

```

Figura 11: Alinhamento das sequências de aminoácidos traduzidas *in silico* dos clones F2 e E3. (*) resíduos idênticos; (:) resíduos altamente conservados; (.) resíduos semi-conservados; (-) marca “gap” na sequência.

As sequências de aminoácidos de F2 e E3 traduzidas *in silico* possuem 77% de identidade. O alinhamento dessas sequências demonstra a presença de uma marca “gap” na posição 190 da sequência dos clones E3 e F2. Entretanto, esse “gap” é um artefato do programa de alinhamento “Clustalw” que posicionou o “stop” códon de E3 erroneamente em relação à sequência de aminoácidos de F2, na busca do melhor alinhamento. Isso significa que a posição 190 no clone F2 é ocupada pelo aminoácido K e que em E3 essa posição corresponde ao “stop” códon. Além disso, se o “stop” códon fosse considerado a rigor, a sequência de aminoácidos de E3 não teria continuidade após esse ponto de parada na tradução. Entretanto, decidiu-se manter a tradução de aminoácidos hipotética a fim de comparar a sequência C-terminal de E3 com a sequência de F2 e de CMS2MS2 extraída do látex de *C. candamarcensis*.

Motivos altamente conservados entre cisteíno-proteases de Caricaceae estão conservados nos clones F2 e E3, como as sequências ₁₁GAVTPV₁₆, ₇₃IKNSWG₁₇₈, além dos resíduos que participam ativamente da catálise enzimática: C₂₅, H₁₅₉ e N₁₇₅. O clone F2 possui as seis cisteínas formadoras de ponte dissulfeto nas posições 22-63, 56-95 e 153-200, importantes para manutenção da estabilidade da

estrutura terciária. Já o clone E3 possui uma substituição da sétima cisteína (C₂₀₀) por uma arginina, além de um “stop” códon na posição 190.

A falta da sétima cisteína, a presença do “stop” códon (figura 11) e dos sítios para as enzimas de restrição Hind III e Xho I (figura 8) impossibilitaram o uso do clone E3 para os experimentos de expressão da proteína, já que a substituição de C₂₀₀ por R₂₀₀ poderiam comprometer a estrutura da proteína recombinante, enquanto que o “stop” códon impediria a expressão da proteína completa. Além disso, a sub-clonagem em vetor pET28a(+)-TEV exige o uso das enzimas Hind III e Xho I no vetor pCR 2.1, impedindo o uso do clone E3 que possui sítios internos para essas enzimas.

O alinhamento da sequência de CMS2MS2 obtida do látex de *C. candamarcensis* com as respectivas sequências traduzidas a partir de F2 e E3 também foi realizado (figura 12).

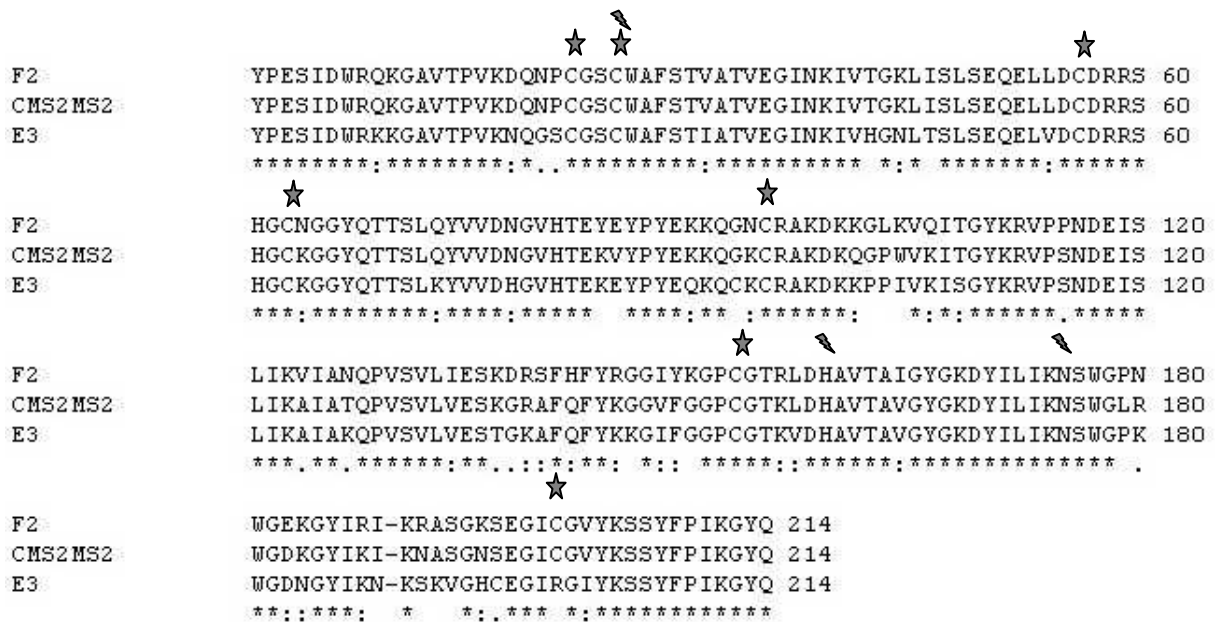


Figura 12: Alinhamento das sequências de aminoácidos obtidas dos clones F2 e E3 com a protease CMS2MS2. (*) resíduos idênticos; (:) resíduos altamente conservados; (.) resíduos semi-conservados; (-) marca “gap” na sequência; (★) marca as cisteínas conservadas; (✂) marca os resíduos da tríade catalítica; (⓪) marca o local do “stop” códon de E3.

A semelhança entre a sequência de aminoácidos de F2 com a de CMS2MS2 foi de 87 %, enquanto que a de E3 com a de CMS2MS2 foi de 83 %. Essa figura possui o mesmo erro de alinhamento causado pelo programa “Clustalw”, como já mencionado para a figura 11.

Também foi feito o alinhamento entre a sequência F2 com a proteína CMS2MS3, outra proteína mitogênica do látex de *C. candamarcensis*, ainda não completamente sequenciada (figura 13).

```

F2          YPESIDWRQKGAVTPVKDQNPCGSCWAFSTVATVEGINKIVTGKLISSLSEQELLDCDRRS 60
CMS2MS3    YPGSVDWRQKGAVTPVGDNPCGSCWAFSTVATVEGINKIVTGHLISSLSEQELLDCDRRS 60
          ** * :*****

F2          HGCNGGYQTTSLSLQYVVDNGVHTEYEYYPYEKKQGNCRACKKGLKVQITGYKRVPPNDEIS 120
CMS2MS3    HGCNGAYQTGSLQYVVDYGVHTEYIYPYE-----RAK-----KVQITGYKRVPTNDEIS 109
          ***** ** ***** ***** ***** ***** ***** ***** *****

F2          LIKVIANQPVSVLIESKDRSFHFYRGGIYKGPCGTRLDHAVTAIGYGKDYILIKNSWGPN 180
CMS2MS3    LIKVIANQPVSVKLIESKGRSFHFYRGGIYKGPCGTRLDHAVTAIGYIRVYILIKNSWGPN 169
          ***** ***** *****

F2          WGEKGYIRIKRASGKSEGICGVYKSSYFPIKGYQ 214
CMS2MS3    WGEKGYIKIKRASGKSEGICGVYKSSYFPTKEYQ 203
          *****:***** * **

```

Figura 13: Alinhamento das sequências de aminoácidos de F2 (traduzida *in silico* a partir da sequência de DNA) e CMS2MS3. (*) resíduos idênticos; (:) resíduos altamente conservados; (.) resíduos semi- conservados nas sequências observadas; (-) marca “gap” na sequência.

A sequência de aminoácidos deduzida do clone F2 possui 91% de identidade com a sequência até então desvendada de CMS2MS3. Já E3 possui uma fraca identidade com CMS2MS3 quando comparado com F2, apresentando 73% de semelhança.

Devido à impossibilidade de usar o clone E3 por motivos mencionados anteriormente, o clone F2 foi escolhido para os experimentos de expressão. Os resultados a partir de agora dizem respeito somente ao clone F2.

O alinhamento da sequência de aminoácidos de F2 com CMS2MS2 (figura 12) demonstra que F2 difere em 27 aminoácidos em relação a CMS2MS2.

A tabela 5 mostra as diferenças entre as sequências de aminoácidos traduzida do clone F2 e da proteína CMS2MS2, tomando como base o alinhamento feito no programa “Clustalw” e utilizando o diagrama de Venn (seção 3.3.12, figura 2) para definir que tipo e natureza da substituição de aminoácidos ocorreu na proteína codificada pelo clone F2.

POSIÇÃO	Proteína		Tipo de substituição	Característica da cadeia lateral
	CMS2MS2	F2 específica		
64	K	N	conservativa	polar
84	K	Y	não-conservativa	-
85	V	E	não-conservativa	-
94	K	N	conservativa	polar
101	Q	K	conservativa	polar
103	P	L	não-conservativa	-
104	W	K	não-conservativa	-
106	K	Q	conservativa	polar
115	S	P	semi-conservativa	pequena
124	A	V	semi-conservativa	pequena
127	T	N	semi-conservativa	pequena
134	V	I	conservativa	alifática
138	G	D	semi-conservativa	pequena
140	A	S	conservativa	muito pequena
142	Q	H	conservativa	polar
145	K	R	conservativa	polar e positiva
164	V	I	conservativa	alifática
179	L	P	não-conservativa	-
180	R	N	conservativa	polar
183	D	E	conservativa	carregada
188	K	R	conservativa	polar e positiva
191	N	R	semi-conservativa	polar
195	N	K	conservativa	polar

Tabela 5: Substituições de aminoácidos da proteína codificada pelo clone F2 em relação a CMS2MS2, mostrando o tipo de substituição e a natureza da cadeia lateral quando a substituição foi conservativa ou semi-conservativa.

A maioria das substituições, 59,3 %, foram conservativas, enquanto que as semi-conservativas e não-conservativas somaram 19,5 e 22,2 %, respectivamente.

4.3 - Sub-clonagem no vetor pET28a(+)-TEV

A sub-clonagem do inserto derivado do clone F2 em vetor pET28a(+)-TEV foi realizada logo após a confirmação da sequência de DNA. Devido a falta de sequências para enzimas de restrição flanqueando as regiões 5' e 3' do fragmento amplificado por PCR, iniciou-se a procura de sítios de restrição disponíveis no vetor pCR 2.1. Esperava-se que, durante a sub-clonagem no vetor de expressão, houvesse a manutenção da fase de leitura e a orientação 5' – 3' do gene a ser inserido. Várias simulações de sub-clonagem foram realizadas, transferindo o gene (derivado da digestão em sítios de restrição de pCR 2.1, figura 1) para inserção nos sítios de restrição dos vetores disponíveis no laboratório. Entretanto, somente a ligação do inserto digerido com HindIII e XhoI (para liberar o inserto do vetor pCR 2.1) com o vetor pET28a(+)-TEV (figura 3) digerido com HindIII e Sall resultou na manutenção da fase de leitura para expressão da proteína recombinante e orientação 5' - 3' do gene. Um esquema representativo do resultado da sub-clonagem está na figura 14.

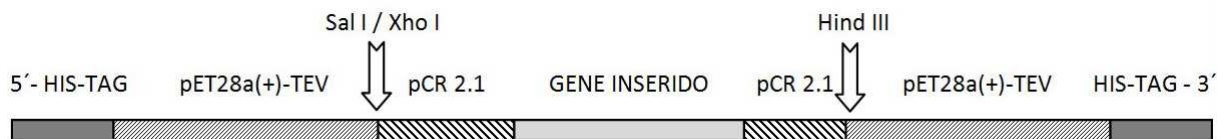


Figura 14: Esquema representativo do resultado da sub-clonagem em pET28a(+)-TEV.

Uma vez escolhidas as enzimas de restrição HindIII e Sall (que possui extremos compatíveis com as geradas por XhoI), o vetor pET28a(+)-TEV foi extraído, purificado após eletroforese em gel de agarose e ligado ao inserto de F2, também previamente digerido e purificado por extração após eletroforese em gel de agarose. O sucesso da etapa de sub-clonagem foi comprovado por PCR utilizando os primers específicos para CMS2MS2, que amplificaram o inserto clonado em pET28a(+)-TEV (figura 15). A fase de leitura e integridade do gene também foram comprovadas por sequenciamento, que mostrou a mesma sequência do clone F2 quando estava no vetor pCR 2.1.

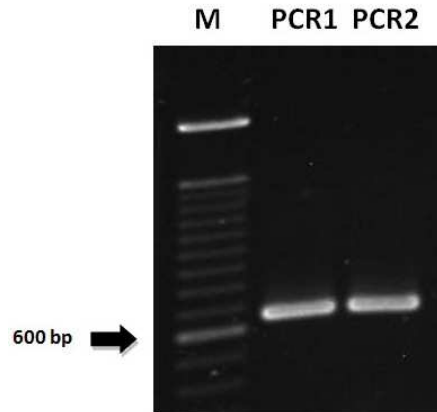


Figura 15: Eletroforese em gel de agarose 1% de fragmento amplificado a partir do inserto F2, subclonado em vetor pET28a(+)-TEV, utilizando os primers específicos para CMS2MS2. Gel de agarose 1% corado com brometo de etídeo. M: padrão de peso molecular 100 bp DNA ladder (Invitrogen). PCR1 e PCR2 são produtos de amplificação de diferentes colônias.

No gel da figura 15 é possível ver a amplificação de um fragmento entre 600 e 700 bp, condizente com o tamanho de 642 bp visualizado nos sequenciamentos.

4.4 - Expressão da proteína recombinante

Uma alíquota contendo a bactéria transformada com o clone que continha a sequência de F2 foi inoculado em meio LB, ao qual posteriormente foi adicionado 0,5 mM IPTG. Essa solução foi incubada em agitador rotatório, a 37°C, com agitação rápida (200 rpm) para expressão da proteína recombinante. Amostras foram coletadas a intervalos de 1 hora, entre 0 e 4 horas. As análises da fração solúvel (figura 16) e insolúvel (figura 17) em géis de poliacrilamida estão representadas abaixo. Dois clones, 1 e 3, foram expressos sob essas condições.

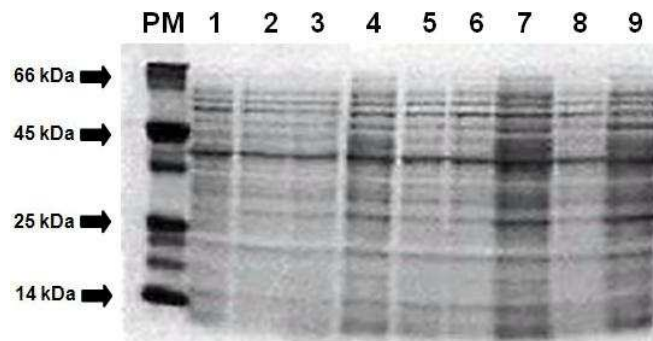


Figura 16: Eletroforese em gel SDS-PAGE da fração solúvel da expressão da proteína recombinante presente no clone 1. PM: padrão de massa molecular composto de BSA (66 kDa), ovoalbumina (45 kDa), α -quimiotripsinogênio(25 kDa) e lisozima (14 kDa). 1: lisado da bactéria BL21-DE3 não-transformada e não-induzida com IPTG; 3, 5, 7 e 9 representam as frações de 0, 1, 2, 3 horas sem adição de 0,5 mM IPTG. 2, 4, 6 e 8 representam as frações de 0, 1, 2 e 3 horas com adição de IPTG.

Na figura 16, as canaletas 2, 4, 6 e 8 mostram o perfil protéico de frações solúveis da expressão do clone que contém a sequência de F2. Esse padrão é bastante semelhante ao apresentado na canaleta 1, que representa a bactéria não-transformada e não-induzida com IPTG. Isso sugere que não houve expressão diferenciada de proteínas de bactérias não-transformadas em relação às de bactérias que continham o vetor pET28a(+)-TEV com o inserto F2.

Em resumo, não há sinal de expressão da proteína recombinante na fração solúvel, cuja massa estimada está ao redor de 30 kDa. Nenhuma banda se sobressaiu sobre as outras, mesmo após 3 horas de indução com IPTG (canaleta 8).

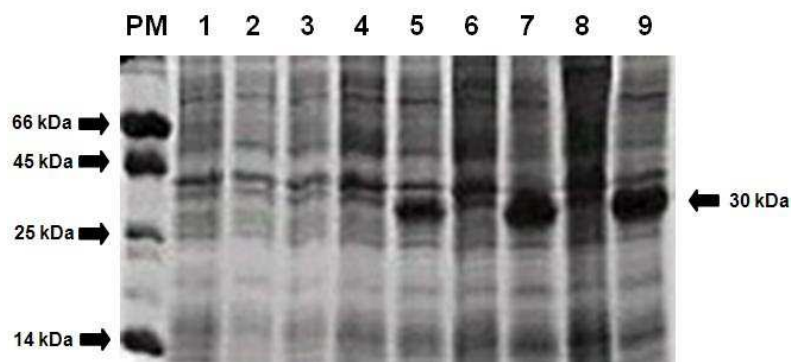


Figura 17: Eletroforese em gel SDS-PAGE da fração insolúvel da expressão da proteína recombinante presente no clone. PM: padrão de massa molecular composto de BSA (66 kDa), ovoalbumina (45 kDa), α -quimiotripsinogênio(25 kDa) e lisozima (14 kDa). 1: lisado da bactéria BL21-DE3 não-transformada e não-induzida. 3, 5, 7 e 9 representam as frações de 0, 1, 2, 3 horas com adição de 0,5 mM IPTG. 2, 4, 6 e 8 representam as frações de 0, 1, 2 e 3 horas sem adição de IPTG.

As proteínas coletadas da fração insolúvel, visualizadas na figura 17, por sua vez, formam um perfil diferente do observado na figura 16. Dessa vez, em relação ao lisado da bactéria não-transformada (canaleta 1) foi possível observar nos tempos de 1, 2 e 3 horas de indução com IPTG (canaletas 5, 7 e 9) a expressão diferenciada de uma proteína de aproximadamente 30 kDa. A banda de 30 kDa, que se sobressai sobre as outras, foi reconhecida como a proteína recombinante e não aparece no lisado da fração insolúvel de bactérias que não foram induzidas com IPTG (canaletas 4, 6 e 8).

A expressão da proteína recombinante, portanto, resultou em uma proteína de cerca de 30 kDa, presente somente na fração insolúvel do lisado bacteriano. A concentração de proteína recombinante é crescente com o aumento do tempo de expressão, como evidenciado nas canaletas 5,7 e 9, que correspondem a 1, 2 e 3 horas de indução da expressão, respectivamente.

A expressão em bactérias que continham somente o vetor vazio pET28a(+)-TEV e a expressão de outro clone que continha o inserto F2 também foram analisadas na presença de IPTG (figura 18).

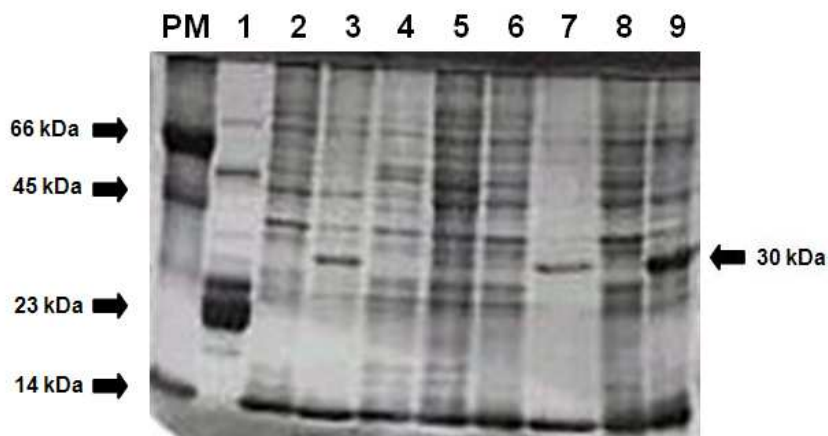


Figura 18: Eletroforese em gel SDS-PAGE da fração insolúvel da expressão da proteína recombinante presente nos clones 1 e 3. PM: padrão de massa molecular composto de BSA (66 kDa), ovoalbumina (45 kDa) e lisozima (14 kDa). 1: P1G10 (23 kDa); 2: clone 3, 1 hora não-induzido; 3: clone 3, 1 hora induzido com 0,5 mM IPTG; 4: pET28a(+)-TEV, 4 horas não induzido; 5: pET28a(+)-TEV, 4 horas induzido com 0,5 mM IPTG; 6: clone 1, 4 horas não-induzido; 7: clone 1, 4 horas induzido com 0,5 mM IPTG; 8: clone 3, 4 horas não-induzido; 9: clone 3, 4 horas induzido com 0,5 mM IPTG.

É possível observar que bactérias que contêm o vetor, mas não o inserto, quando induzidas com 0,5 mM IPTG (canaletas 4 e 5) não apresentam a expressão da proteína recombinante na fração insolúvel, como visto nas canaletas 3, 7 e 9.

4.5 - Análise dos parâmetros bioquímicos



A partir dos sequenciamentos dos clones 1 e 3, que contêm o inserto F2 subclonado em pET28a(+)-TEV, foi possível delimitar as sequências de DNA que correspondiam aos vetores pCR 2.1 e pET28a(+)-TEV, unidas após a sub-clonagem. Elas foram verificadas em vários sequenciamentos de DNA, a fim de confirmar a autenticidade das sequências fornecidas pelo sequenciador automático.

Uma vez feita a confirmação, essas sequências de DNA foram traduzidas *in silico* com a ferramenta “Translate tools” presente no servidor Expasy. A construção final da sequência putativa de CMS2MS2 contendo as partes dos vetores pCR 2.1 e pET28a(+)-TEV, resultou numa proteína de 292 aminoácidos, cuja sequência deduzida está representada na figura 19. Nesta sequência foram observadas as mesmas variações vistas no sequenciamento do inserto em pCR 2.1.

```

HHHHHHENLYFQGHMASMTGGQQMGRGSEFELRRAAASVMDICR
RLY PESIDWRQKGA VTPVKDQNPCGSCWAFSTVATVEGINKIVTG
KLISLSEQELLDCDRRSHGCNGGYQTSLQYVVDNGVHTEYEYPEYK
KQGNCRAKDKKGLKVQITGYKRVPPNDEISLIKVIANQPVSVLIESK
DRSFHFYRGGIYKGPCGTRLDHAVTAIGYGKDYILIKNSWGPNWGE
KGYIRIKRASGKSEGICGVYKSSYFPIKGY QKPNSSTLAAVTSGSELG
TKLAAALEHHHHHH STOP

```

Figura 19: Sequência de aminoácidos traduzida *in silico* do clone F2, mostrando as duas caudas de histidina, o “stop” códon, e em preto o início e final da sequência codificante de CMS2MS2. O tracejado  representa as partes da sequência advindas do vetor pCR 2.1 e  indica os aminoácidos advindos do vetor pET 28a(+)-TEV.

Em preto, os aminoácidos Y e Q representam os resíduos N e C-terminal da proteína recombinante. Além disso, ela possui duas caudas de histidina, sendo que o “stop” códon aparece logo após a cauda de histidina C-terminal.

O ponto isoelétrico (pI) e a massa molecular (MW) das proteínas foram preditos pelo programa ScanSite, utilizando a porção específica de F2

(F2 específica). O pI e MW da proteína F2 também foram calculados acrescentando-se as partes não específicas de CMS2MS2 (F2 total).

Proteína	pI	MW (Da)	Número de aminoácidos
CMS2MS2	9.31	23590.76	214
F2 específica	9.18	23944.03	214
F2 total	9.18	32603.75	292

Tabela 6: Parâmetros bioquímicos de CMS2MS2 e da proteína recombinante F2.

O resultado mostrou que os valores teóricos de pI e MW de CMS2MS2 e F2 específica são muito semelhantes. Além disso, o acréscimo de regiões dos vetores pCR 2.1 e PET28a(+)-TEV durante a sub-clonagem não alteraram o valor de pI da proteína recombinante total.

4.6 - Purificação da proteína recombinante

Utilizando o clone F2, uma nova expressão da proteína recombinante foi realizada utilizando 4 horas de indução, tempo escolhido para a expressão após os testes iniciais. Novas coletas de lisado bacteriano foram realizadas para isolamento de cerca de 4 mg de proteína da fração insolúvel. Essa é a quantidade ideal para a purificação da banda de 30 kDa com o menor número de bandas retidas na coluna de afinidade que advinham da bactéria. A quantidade ideal havia sido determinada previamente realizando-se purificações-teste em pequena escala e verificação do rendimento em SDS-PAGE.

O pellet foi dissolvido diretamente no tampão de ligação, contendo uréia 8M, e aplicado à coluna de afinidade ao níquel. Após lavagens sucessivas com o tampão de ligação para eluição das proteínas que possuíam pouca afinidade com a coluna, o tampão de eluição contendo imidazol a 250 mM foi aplicado e a proteína recombinante foi liberada da coluna, uma vez que o imidazol, presente em altas concentrações no tampão de eluição, compete com a cauda de histidina pelo níquel, liberando a proteína da resina.

As frações resultantes da cromatografia foram submetidas à gel desnaturante, como visto na figura 20.

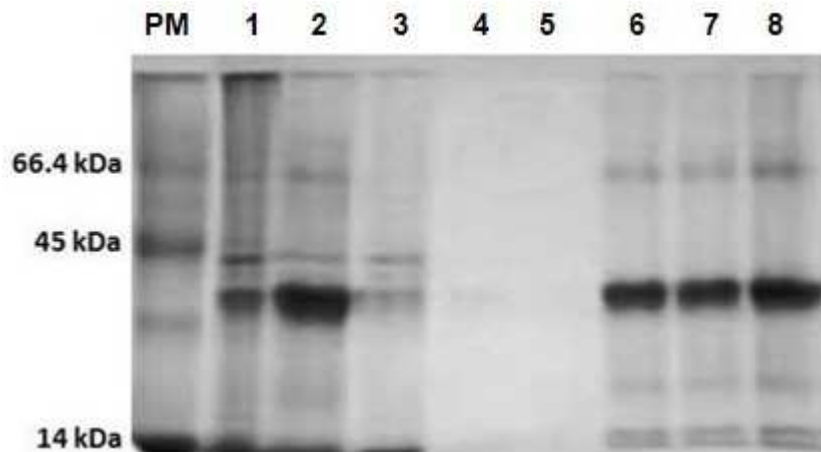


Figura 20: Eletroforese em SDS-PAGE mostrando as frações eluídas da coluna de afinidade. PM: BSA (66.4 kDa), ovoalbumina (45 kDa) e lisozima (14 kDa). 1 e 2: frações insolúveis de expressão do clone 1, induzido com 0,5 mM IPTG; 3, 4 e 5: frações coletadas de sucessivas lavagens com tampão de ligação; 6, 7 e 8: frações coletadas após aplicação do tampão de eluição na coluna de afinidade, mostrando a banda de 30 kDa correspondente à proteína recombinante.

Após a purificação em cromatografia de afinidade foi possível estimar a recuperação de 35 mg/L de cultura bacteriana. É possível purificar cerca de 2 mg de proteína recombinante em cada passagem pela coluna de afinidade ao níquel, com dimensões de 13 cm de altura por 0,5 cm de diâmetro.

4.7 - Ensaios de atividade biológica com CMS2MS2 recombinante

De posse da proteína recombinante parcialmente purificada, iniciaram-se os testes de atividade biológica. A medida da atividade mitogênica de CMS2MS2 recombinante foi realizada através do ensaio de incorporação de [H^3 -metil]-timidina no DNA, como descrito em 3.3.19.

Foi demonstrado previamente que a proteína CMS2MS2 extraída do látex de *Carica candamarcensis* possui atividade mitogênica sobre fibroblastos L929 (Gomes *et al.*, 2005) e em osteoblastos obtidos de cultura primária (dados não publicados). Considerando que a proteína recombinante se encontrava em um tampão rico em uréia (8M), era previsível que esta estivesse em sua forma desnaturada. Entretanto, como foi demonstrado que a atividade proliferativa de CMS2MS2 extraída do látex não depende de sua atividade proteolítica (Gomes, 2008), testou-se a proteína

recombinante dissolvida em tampão de eluição (seção 3.1.2), esperando-se que determinantes conformacionais não fossem essenciais para a atividade mitogênica.

Neste ensaio, as atividades proliferativas de CMS2MS2 obtida do látex de *C. candamarcensis* e também da proteína recombinante (advinda do clone F2) foram analisadas. As células escolhidas foram fibroblastos imortalizados da linhagem L929 e osteoblastos obtidos de cultura primária extraídos de calvária de ratos Wistar. Foram utilizadas concentrações da proteína recombinante semelhantes às concentrações de CMS2MS2 não-recombinante que tinham comprovado efeito mitogênico. Os resultados dos experimentos com fibroblastos L929 e osteoblastos de cultura primária são mostrados nas figuras 21 e 22, respectivamente.

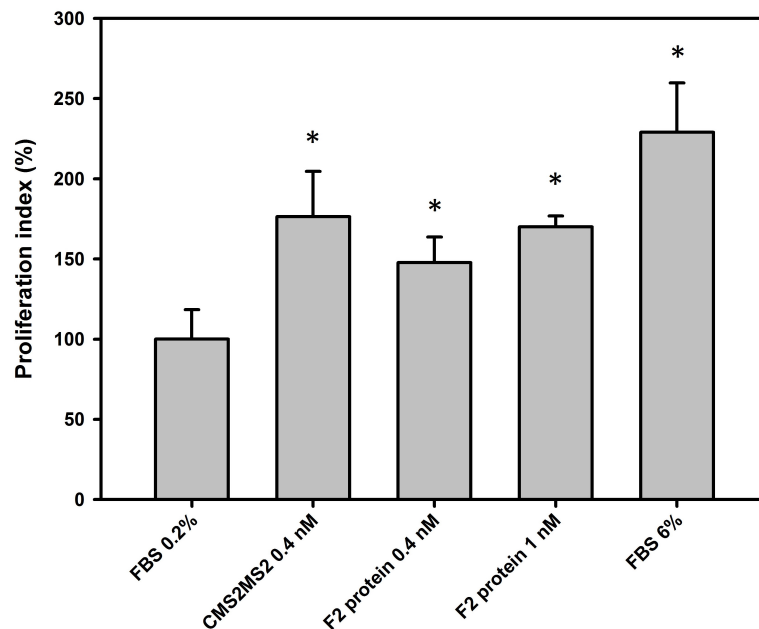


Figura 21: Índice de proliferação de células L929 após tratamento com CMS2MS2 não-recombinante e proteína recombinante. A proliferação foi expressa como “índice de proliferação”, atribuindo um valor de 100% à proliferação celular obtida com FBS 0,2 %, condição de careciamento. (*) é a diferença significativa em relação ao controle (FBS 0,2%), com $p < 0,05$ no teste “t de Student”. As barras representam a média \pm desvio padrão, com “n” de poços analisados igual a 4..

A atividade mitogênica em fibroblastos é 47% maior do que no controle FBS 0,2% utilizando 0,4 nM de proteína recombinante e 70% usando 1 nM de proteína.

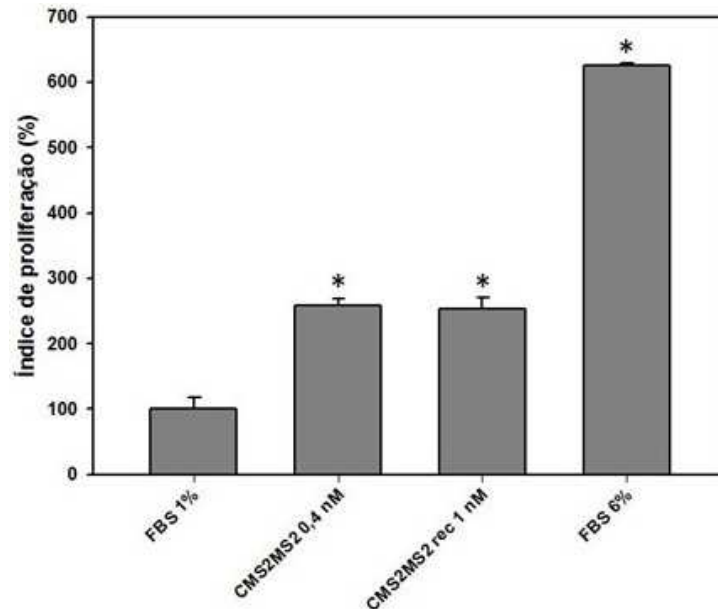


Figura 22: Índice de proliferação de osteoblastos após tratamento com CMS2MS2 natural e CMS2MS2 recombinante. A proliferação foi expressa como “índice de proliferação”, atribuindo um valor de 100% à proliferação celular obtida com FBS 1%. (*) é a diferença significativa em relação ao controle (FBS 1%), com $p < 0,05$ no teste “t de Student”. As barras representam a média \pm desvio padrão, com “n” de poços analisados igual a 2.

Em osteoblastos obtidos de cultura primária a indução da mitose por 1 nM de proteína recombinante é 152% maior do que no controle negativo FBS 1%, que é a condição de carenciamento para esse tipo celular.

Apesar de ter sido demonstrado que a atividade de CMS2MS2 obtida a partir do látex apresenta atividade entre 0,4 nM e 1 nM, testes com outras concentrações da proteína recombinante não foram realizados. Os experimentos mostrados acima foram realizados utilizando o mesmo estoque de uma purificação em coluna de afinidade ao níquel. Anterior a esses, outros experimentos foram realizados com a banda de 30 kDa eletroeluída de gel de poliácridamida 12%. Nesses ensaios, também se observou atividade proliferativa induzida pela proteína recombinante, descartando a possibilidade de que as proteínas de bactérias contaminantes advindas da purificação em coluna de afinidade ao níquel pudessem ser as

responsáveis pela atividade mitogênica observada nas figuras 21 e 22 (resultados não mostrados).

Em resumo, foi possível observar que nas concentrações de 0,4 e 1 nM a proteína recombinante apresenta atividade mitogênica sobre fibroblastos e a 1 nM foi capaz de estimular a proliferação em osteoblastos obtidos de cultura primária. Esses dados sugerem que determinantes conformacionais podem ser pouco importantes para a indução proliferativa causada pelas proteases de *C. candamarcensis* que até então foram testadas quanto à essa capacidade.

5. Discussão

Em estudos anteriores, havia sido demonstrado que a protease CMS2MS2 do látex de *Carica candamarcensis* apresenta atividade mitogênica sobre fibroblastos L929 (Gomes et al., 2005). Um dos objetivos mais importantes do nosso grupo devido à apresentação dessa atividade biológica *in vitro* é o estudo *in vivo* dos efeitos da protease mitogênica em modelos de cicatrização, que demandam grandes quantidades de proteína. Porém, a elaborada obtenção de CMS2MS2 a partir do látex por processos cromatográficos é um processo trabalhoso e demorado, com rendimentos modestos da protease mitogênica. Diante do exposto, o objetivo desse trabalho foi produzir CMS2MS2 recombinante para posteriores estudos da atividade biológica e caracterização estrutural da proteína.

A obtenção de fragmentos de 500 e 650 bp por PCR (figura 5) foi realizada utilizando um par de primers degenerados baseados na sequência N e C-terminal da proteína, recentemente elucidada por Gomes et al. (2007). A clonagem do fragmento de 650 bp foi realizada em vetor pCR 2.1 e os clones positivos foram selecionados através de PCR, digestão com enzimas de restrição (figura 6) e sequenciamento (figuras 7 e 8).

O inserto oriundo do clone F2 foi escolhido para a expressão e sub-clonado no vetor pET28a(+)-TEV. O perfil de expressão de clones positivos em bactérias BL21-DE3 induzidas com IPTG demonstrou uma proteína de aproximadamente 30 kDa (figura 17), como esperado após a construção no vetor de expressão. A proteína recombinante foi encontrada somente na fração insolúvel do lisado bacteriano. A purificação da proteína recombinante (figura 20) e sua atividade biológica mitogênica em fibroblastos L929 e osteoblastos de cultura primária foram demonstradas (figuras 21 e 22).

Com a utilização dos primers “forward” e “reverse”, cujas sequências tinham como base as sequências N e C-terminal da proteína, era esperada a amplificação do mRNA correspondente à cisteíno-protease CMS2MS2. O fragmento de 650 bp amplificado por PCR, que possuía tamanho próximo ao esperado para o tamanho do gene da proteína CMS2MS2, foi então, utilizado para clonagem em vetor pCR 2.1.

O sequenciamento dos clones revelou que os fragmentos originados da PCR possuíam 642 nucleotídeos e codificavam para cisteíno-proteases semelhantes, mas

não idênticas, a CMS2MS2. Dois dos fragmentos presentes nesses clones, E3 e F2, foram objetos do estudo.

Por ter sido usado um par de primers degenerados, que contêm um conjunto de várias sequências N e C-terminal possíveis, foi observada a amplificação de possíveis isoformas de cisteíno-proteases, semelhantes principalmente a CMS2MS2 e CMS2MS3 (figuras 12 e 13), além daquelas que não tiveram suas sequências de DNA extensivamente analisadas devido à deficiente qualidade do sinal do sequenciamento.

A semelhança de sequências N-terminais entre proteínas de *C. candamarcensis* que se originam de um mesma fração cromatográfica foi descrita por Teixeira et al. (2008). Nesse trabalho, fica evidente que muitas proteases dessa espécie possuem as mesmas ou muito semelhantes sequências N-terminais. Entretanto, dados de espectrometria de massas mostrou que as proteínas completas possuem tamanhos diferentes, mesmo tendo sequências N-terminais idênticas (Teixeira et al., 2008). Esses dados corroboram para a hipótese da ocorrência de isoformas em *C. candamarcensis*.

Várias isoformas já haviam sido descritas para a quimopapaína, de *C. papaya* (Taylor et al., 1999) e para proteases das espécies *V. stipulata* e *V.x heilbornii* (Kyndt et al., 2007). O sequenciamento de clones derivados de uma biblioteca de cDNA de folhas de *C. papaya* identificou cinco cDNAs codificando para quimopapaínas em sua forma precursora (Taylor et al., 1999). Também através de bibliotecas de cDNA, foram descobertas isoformas de cisteíno-proteases de *V. stipulata* e *V. x. heilbornii* utilizando primers que se anelavam a regiões conservadas de cisteíno-proteases (Kyndt et al., 2007).

As diferenças nas sequências nucleotídicas dos clones F2 e E3, e desses com a de CMS2MS2, também podem resultar do uso de uma *Taq* polimerase não “proof-reading”. Essa *Taq* polimerase aumenta a chance de substituições errôneas, de maneira aleatória. Porém a taxa de erro da *Taq* polimerase (entre 1×10^{-4} a 2×10^{-5} por par de base) não é suficiente para explicar a magnitude de variação observada entre as sequências de aminoácidos traduzidas *in silico* de F2 e E3 (49 aminoácidos) ou então entre as sequências de F2 e CMS2MS2 (27 aminoácidos). A variação esperada estaria em torno de 6×10^{-2} a 3×10^{-3} erros por amplicom em cada ciclo.

Além disso, o uso de temperatura de anelamento de 39°C (ver seção 3.3.3) pode ter diminuído as condições de estringência da PCR a ponto de possibilitar a amplificação de mais isoformas de cisteíno-proteases, o que poderia não acontecer caso a temperatura de anelamento fosse elevada. Entretanto, a temperatura para a amplificação com os primers degenerados teve que ser drasticamente diminuída para obtenção das bandas correspondentes à cisteíno-proteases. Outro fator inerente à reação de PCR que poderia aumentar os erros durante a amplificação é o elevado número de ciclos da reação padronizada, que aumenta de acúmulo dos erros da *Taq* polimerase de forma exponencial (Eckert & Kunkel, 1991).

É provável que uma combinação de fatores incluindo os erros induzidos pela *Taq* polimerase, as condições de reação de PCR com a ocorrência de isoformas no cDNA sejam responsáveis pelas substituições de aminoácidos observadas.

É possível que o conjunto de mRNAs presente na folha de *C. candamarcensis* codifique várias isoformas de cisteíno-proteases, que podem exercer funções semelhantes, mas ligeiramente sub-especializadas nas células. A expressão de isoformas explicaria a diversidade de clones obtidos durante a amplificação por CMS2MS2, exemplificada pelos clones E3 e F2.

Assim como foi sugerido para o surgimento de isoformas de quimopapaína de *C. papaya* (Taylor et al., 1999) e também para as cisteíno-proteases do gênero *Vasconcellea* (Kyndt et al., 2007), duplicações gênicas e sub-especializações de cisteíno-proteases ligeiramente diferentes em sua sequência de aminoácidos também poderiam ter ocorrido na espécie *C. candamarcensis*. A construção de uma biblioteca de cDNA composta de sequências de cisteíno-proteases de *C. candamarcensis*, como realizado por Kyndt e colaboradores (2007) para *V. stipulata* e *V. x heilbornii*, permitira a observação e quantificação de isoformas dessas proteínas em *C. candamarcensis*.

A obtenção das sequências de nucleotídeos de proteases de *C. candamarcensis* aumentou o número de sequências de DNA descritas para as cisteíno-proteases dessa espécie. Dentre elas, E3 e F2 tiveram suas sequências de DNA e aminoácidos traduzidas *in silico* analisadas pelas ferramentas “blastn”, “blastx”, Clustalw e NETNGLYC (seção 3.3.12).

Comparando a sequência de nucleotídeos de F2 e E3 com outras cisteíno-proteases presentes no banco de dados do NCBI foi possível observar uma alta identidade com sequências de proteases de *V. stipulata* e *V. x heilbornii* (tabelas 1 e

2). Essa semelhança já havia sido constatada por Kyndt e colaboradores (2007), cujo grupo tem mostrado evidências de que a espécie *V. x heilbornii* pode ter surgido como um híbrido de *C. candamarcensis* e *V. stipulata*. A alta identidade entre estas sequências de DNA sugere que essa hibridação ocorreu recentemente e que não houve tempo suficiente para grandes mudanças no material genético dessas espécies. Novamente, essa hipótese só poderá ser confirmada realizando-se um estudo mais abrangente, como a construção de bibliotecas de cDNA dessas espécies que também contenham outros genes não relacionados à cisteíno-proteases.

É interessante observar também que a sequência de DNA do clone F2 possui mais semelhança com as sequências gênicas de proteases da família *Vasconcellea* do que E3, que possui alta identidade com isoformas de quimopapaína (tabelas 1 e 2). Essa relação sugere que *C. candamarcensis* possa conter uma mistura de mRNA codificando proteases advindos de hibridação das espécies de *Vasconcellea* com *C. papaya*.

A sequência de DNA de F2 possui maior semelhança com CC23, de *C. candamarcensis* (94%), do que com CMS2MS2 (87%). Entretanto, a proximidade entre F2 e CMS2MS2 é maior se comparada à sequência obtida pela tentativa de amplificação de CMS2MS2 com outro par de primers (número de acesso AY194124) por Junqueira (2007), que apresentou 60 % de identidade com CMS2MS2.

Examinando a sequência de aminoácidos de E3, traduzida *in silico* a partir da sequência de DNA, foi possível observar a presença de dois sítios potenciais de glicosilação. Um desses sítios, ${}_{44}\text{NLT}_{46}$ já havia sido observado na proteína CC-III (Jaziri et al, 1994), indicando pela segunda vez a possibilidade de uma cisteíno-protease de *Carica candamarcensis* com sítio de glicosilação em sua porção madura. É possível que E3 seja uma isoforma de CC-III, uma vez que apresenta 92% de identidade com esta última. A presença do sítio de glicosilação pode não ter surgido pelos erros da *Taq* polimerase, já que CC-III, uma proteína extraída da mesma espécie, *C. candamarcensis*, possui esse sítio (Jaziri et al., 1994). O motivo ${}_{44}\text{NLT}_{46}$ não é detectado em CMS2MS2, sustentando a possibilidade que E3 e CMS2MS2 possam ser produtos de genes diferentes.

A sequência do clone E3 conservou os resíduos C₂₅, H₁₅₉ e N₁₇₅, responsáveis pela catálise enzimática. Entretanto, a sequência desse clone mostra a substituição da sétima cisteína por uma arginina (figura 12). A falta dessa cisteína,

importante na formação de uma ponte dissulfeto, poderia comprometer o enovelamento da proteína codificada por E3, caso esse clone fosse expresso. Entretanto, seria interessante averiguar se a falta da porção C-terminal da proteína que seria codificada pelo clone E3 é essencial para a atividade mitogênica. Isso poderia ser feito utilizando o sítio de XhoI e o segundo sítio de Hind III (que se apresenta depois da seqüência do primer “reverse” na figura 8). O uso das enzimas de restrição nessas regiões e sub-clonagem no vetor pET28a(+)-TEV levaria a produção de uma proteína com 189 aminoácidos, ao invés de 214 aminoácidos. A falta dos 25 resíduos C-terminais poderia oferecer indícios se essa seqüência é importante para a atividade mitogênica.

A expressão de E3 também não foi considerada devido a presença de um “stop” códon e um sítio de clivagem para a enzima de restrição HindIII no interior da seqüência codificante da proteína (figura 8), o que tornaria improvável a chance de sucesso da sub-clonagem. O “stop” códon pode ser fruto do uso da *Taq* polimerase não “proof-reading”, mas não se sabe se o sítio de HindIII ocorre na seqüência natural dessa isoforma ou se existe devido ao erro aleatório da *Taq* polimerase.

Também existe a possibilidade que o clone E3 corresponda à seqüência de um pseudogene. Considerando a hipótese de que houve duplicações gênicas nas espécies do gênero *Vasconcellea*, é provável que esses eventos levaram a duplicação de uma cisteíno-protease que, modificada com um “stop” códon, se tornou não-funcional.

Já F2, o outro clone analisado, não possui o sítio de glicosilação observado em E3, e maior identidade com CMS2MS2 do que a seqüência do clone E3. Em F2 a tríade catalítica composta por cisteína 25, histidina 159 e asparagina 175 foi preservada em sua posição original, assim como as pontes dissulfeto entre as cisteínas 22-63, 56-95 e 153-200, importantes para manutenção da estabilidade da estrutura terciária (figura 12). Outras regiões altamente conservadas em cisteíno-proteases provenientes de *Caricaceae*, como por exemplo, ¹¹GAVTPV₁₆ e ¹⁷³IKNSWG₁₇₈, também estão presentes em F2 e E3 (figura 12).

Assim como CMS2MS2 obtida do látex de *C. candamarcensis*, o clone F2 não contém a seqüência de aminoácidos responsável pelo “extra-loop” observado entre as posições 168-171 de quimopapaína, caricaína ou glicil-endopeptidase de *C. papaya*. A ausência dessa seqüência aproxima filogeneticamente a proteína produzida pelo clone F2 com outras proteínas de *C. candamarcensis*, *V. stipulata* e

V. x heilbornii, além de papaína de *C. papaya*, que também não possuem o “extra-loop” (Kyndt et al., 2007).

Ao comparar as sequências de aminoácidos de F2 e E3 com sequências de proteínas depositadas no banco de dados foi possível observar novamente identidade com cisteíno-proteases da família Caricaceae. A sequência de aminoácidos traduzida *in silico* de E3, por exemplo, possui alta identidade com a de CC-III, uma enzima de *C. candamarcensis* cuja sequência já era conhecida (tabela 4). Isso já era esperado, já que CC-III havia apresentado semelhança com E3 devido ao sítio de glicosilação. A sequência protéica de F2, por sua vez, codifica uma proteína muito semelhante a VS-B, de *V. stipulata* (tabela 3).

As estruturas primárias putativas de E3 e F2 também possuem alta identidade com outras cisteíno-proteases. As duas sequências são muito parecidas com mexicana (F2:83%, E3:73% de identidade) e quimomexicana (F2:84%, E3:74% de identidade), que são cisteíno-proteases da espécie *Jacaratia mexicana*, que também pertence à família Caricaceae (tabelas 3 e 4). A proximidade evolutiva das proteases de *C. candamarcensis*, especialmente de CMS2MS2, com as proteínas de *J. mexicana* já havia sido observada por Gomes et al. (2007).

Além disso, foi observado que a sequência traduzida de F2 possui mais identidade com CMS2MS3 (91 %) (figura 13) do que CMS2MS2 (87%). Foi demonstrado que tanto CMS2MS2 quanto CMS2MS3 possuem atividade mitogênica sobre fibroblastos L929 (Gomes et al., 2005; Gomes, 2008) apesar de existirem diferenças na estrutura primária entre essas proteases. Isso indica que a proteína codificada pelo clone F2, CMS2MS2 e CMS2MS3 possuem pontos em comum que determinam que essas, e não outras proteases de *Carica candamarcensis*, sejam mitogênicas.

Um outro ponto a ser observado é que a sequência N-terminal de CMS2MS3 apresenta uma glicina ao invés de uma serina na posição 3 e uma valina ao invés de uma isoleucina na posição 5 (figura 13). Isso indica que o gene amplificado pelo clone F2 não é o de CMS2MS3, mas se parece muito com essa proteína. Isso mais uma vez indica que o uso de primers degenerados deve ter levado à amplificação de proteínas muito semelhantes à CMS2MS2 e CMS2MS3, mas não idênticas com aquelas obtidas do látex da planta, cuja estrutura primária foi determinada pela degradação de Edman. O uso de primers não-degenerados, poderá diminuir o número de cDNAs que são amplificados e a análise das isoformas se tornará mais

simples. Um par de primers degenerados foi desenhado para amplificação da proteína com o pro-peptídeo e outro par de primers não-degenerados já foi desenhado para amplificação da proteína madura recombinante sem extensões para regiões do vetor. O sequenciamento dos amplicons utilizando esses primers direcionará a análise das isoformas de cisteíno-proteases, especialmente as de CMS2MS2.

Os parâmetros bioquímicos analisados no programa ScanSite mostraram que a parte específica da proteína recombinante derivada do clone F2 possui um pI semelhante ao de CMS2MS2. Esse fato mostra que a proteína recombinante possui distribuição total de cargas semelhante à de CMS2MS2 e que essa propriedade, pelo menos, não seria limitante para a proteína recombinante apresentar as mesmas funções que CMS2MS2 natural.

As diferenças encontradas entre as sequências de F2 e CMS2MS2 não impediram a tentativa de expressão da proteína recombinante codificada pelo clone F2. Entretanto, durante a sub-clonagem no vetor pET28a(+)-TEV houve o acréscimo de 78 aminoácidos na proteína recombinante, derivados do próprio vetor pET28a(+)-TEV e também do vetor anterior, pCR 2.1 (figura 19). Mesmo assim, pretendeu-se expressar e verificar a atividade biológica da proteína recombinante.

A análise do lisado bacteriano derivado da expressão da proteína recombinante não a mostrou na parte solúvel (figura 16). Essa constatação indica que a proteína provavelmente está sendo expressa e levada a corpos de inclusão no citoplasma da bactéria (Bromme et al., 2004). A presença da proteína na fração insolúvel foi detectada, como previsto (figura 17). Entretanto, devido à alta condição desnaturante a qual a proteína foi submetida durante a solubilização dos corpos de inclusão, provavelmente a proteína recombinante pode não possuir a mesma conformação que a proteína CMS2MS2 natural apresenta.

Após a purificação obteve-se uma proteína recombinante quase pura, contaminada por algumas proteínas de bactérias que também possuem forte afinidade ao níquel (Henge, 1995). Entretanto, a quantidade de proteína recombinante sempre representava mais do que aproximadamente 90 % do total de proteínas, como apreciado no gel de poliacrilamida da figura 20. Pela análise em gel também era possível ver que o rendimento da proteína recombinante após a purificação estava em 35 mg/L de cultura bacteriana.

Esse rendimento é superior à muitas expressões e purificações de cisteíno-proteases semelhantes à papaína (Bromme et al., 2004), cuja média é de 20 mg/ml de proteína. Foi descrito para a expressão de quimopapaína em *E. coli* rendimentos de 4 a 6 mg por litro de cultura bacteriana (Taylor et al., 1999). Já a expressão de papaína em *E. coli* resultou em 3mg/L de expressão (Taylor et al., 1992), em baculovirus 300 µg/ L (Vernet et al., 1991) e em levedura a propapaína apresentou um rendimento de 1.7 mg/L (Ramjee et al., 1996). Recentemente a expressão de papaína em *E. coli* produziu o rendimento de 400 mg/L de cultura bacteriana (D. Choudhury, comunicação pessoal). Desse modo, o rendimento da expressão da proteína recombinante descrita nesse trabalho está de acordo e chega ser superior aos rendimentos descritos na literatura.

Testes de atividade proteolítica com substratos comumente usados para cisteíno-proteases foram feitos, mas a proteína recombinante purificada não apresentou atividade proteolítica, mesmo quando testada em condições ótimas de tampão e substrato. A falta de atividade proteolítica pode ser devido ao fato da proteína estar desenovelada, já que se encontra em uréia 8M.

Tentativas de diálise da solução contendo a proteína recombinante foram realizadas. Entretanto, havia perda significativa de material durante a diálise. Esse fato pode ser devido à agregação das proteínas na membrana de diálise ou também devido a um possível enovelamento, ativação e auto-proteólise da proteína recombinante. Entretanto, a primeira hipótese é mais provável, já que pequenos precipitados se formam na membrana durante a diálise.

Outros dois fatores podem estar dificultando o enovelamento da proteína recombinante. Foi mostrado que a porção pro-peptídica é essencial ao enovelamento de muitas cisteíno-proteases da família Caricaceae (Bromme et al., 2004). Portanto, não se sabe se a falta da sequência pro-peptídica inviabilizaria o enovelamento da proteína recombinante mesmo em condições de re-enovelamento. Outro fator seria a presença de regiões não codificantes de CMS2MS2, ou seja, as regiões de pCR2.1 e pET28a(+)-TEV presentes na proteína codificada pelo clone F2, que também poderiam inviabilizar o enovelamento da proteína recombinante em sua forma proteoliticamente ativa. Além disso, testes de re-enovelamento ainda não foram realizados para essa proteína, pois a atividade proteolítica poderia não ser importante para a atividade mitogênica, principal alvo de nosso estudo, já que foi

demonstrado que para CMS2MS2 obtida do látex essa propriedade não é essencial para a indução proliferativa de fibroblastos L929 (Gomes, 2008).

A expressão da proteína CMS2MS2 utilizando uma *Taq* polimerase “proof-reading” e sem acréscimos de extensões é essencial para verificar se é possível expressar a proteína em sistema heterólogo em sua forma proteoliticamente ativa, sem sua porção pro-peptídica. A expressão da proteína CMS2MS2 em sua forma precursora já está sendo realizada, mas ainda não possuímos dados se essa porção é essencial ao envelhecimento ou à atividade proteolítica, pois experimentos de autólise do pro-peptídeo e atividade proteolítica da proteína sem essa porção ainda não foram realizados.

Apesar das diferenças observadas e já discutidas entre o produto de expressão do clone F2 e CMS2MS2 derivada da planta, a atividade mitogênica da proteína recombinante parece não ter sido comprometida (figuras 21 e 22).

Gomes (2008) já havia demonstrado que a atividade mitogênica de CMS2MS2 independe de sua atividade proteolítica. Observando esse resultado, esperávamos que a proteína recombinante, mesmo não tendo a conformação apropriada, pudesse apresentar atividade mitogênica, já que determinantes lineares, e não conformacionais, poderiam ser os responsáveis pela atividade mitogênica de CMS2MS2. Os resultados de experimentos de atividade mitogênica com a proteína recombinante (figuras 21 e 22) reforçam essa hipótese.

A proteína recombinante codificada pela sequência do clone F2 apresentou atividade proliferativa sobre fibroblastos L929 nas concentrações de 0,4 e 1 nM, e em osteoblastos obtidos de cultura primária na concentração de 1 nM. Essas concentrações equivalem àquelas usadas nos experimentos de atividade mitogênica com a proteína natural (de 0,4 a 2 nM). O índice de proliferação (em relação ao controle basal FBS 0,2%) que a proteína recombinante exibiu sobre fibroblastos L929 é de 47 e 70% a 0,4 nM e 1 nM de proteína recombinante, respectivamente, e 152 % em osteoblastos de cultura primária. Esse índice é compatível com a ação de 0,4 nM de CMS2MS2 natural, que resultou em índices de proliferação de 68% para fibroblastos L929 e 158% para osteoblastos de cultura primária.

Não foi averiguado, porém, como a proteína recombinante estaria estimulando a mitogênese. Em relação à proteína CMS2MS2 natural alguns dados já foram comprovados. Sabe-se que a atividade de CMS2MS2 do látex depende da ativação da via das MAP-quinases, e que muito provavelmente a ação de CMS2MS2

resulta na liberação de fatores de crescimento pelas células, como evidenciado em testes utilizando meio condicionado coletado algumas horas após a aplicação de CMS2MS2 (Gomes, 2008). Desse modo, testes de inibição da via das MAP-quinases e verificação da atividade mitogênica são essenciais para saber se a proteína recombinante age pela mesma via de sinalização que CMS2MS2 extraída do látex de *C. candamarcensis*. Além disso, o uso de meio condicionado após a aplicação da proteína recombinante seria adequado para verificar se a proteína recombinante também poderia induzir a liberação de fatores de crescimento pelas células.

A verificação de que a proteína codificada pelo clone F2 age da mesma maneira que CMS2MS2 é importante para identificar, principalmente, quais determinantes, lineares ou conformacionais, são importantes na resposta mitogênica de fibroblastos L929 e osteoblastos de cultura primária a estas proteínas. A clonagem do gene para CMS2MS2, ou isoformas desta, utilizando *Taq* polimerase “proof-reading” também seria muito importante para determinar se somente CMS2MS2 e CMS2MS3 do látex possuem essa atividade mitogênica, ou se pequenas variações na sequência torna outras isoformas candidatas a ter essa atividade.

6. Conclusões

Através de PCR do cDNA, obtido do RNA de folhas de *Carica candamarcensis*, foi possível amplificar fragmentos de aproximadamente 650 e 500 bp utilizando os primers CMS2MS2 “forward” e “reverse”.

O sequenciamento do fragmento de aproximadamente 650 bp, clonado em vetor pCR 2.1, mostrou seqüências codificadoras para cisteíno-proteases. As seqüências de DNA dos clones F2 e E3, são semelhantes às de outras cisteíno-proteases de *Carica papaya* e de espécies do gênero *Vasconcellea*.

A tradução *in silico* de aminoácidos de F2 e E3 e seu alinhamento com outras seqüências de proteases demonstrou que essas proteínas não só se parecem com cisteíno-proteases de *Carica papaya* e de espécies do gênero *Vasconcellea*, mas também com cisteíno-proteases de *Jacaratia mexicana*, espécie que também pertence à família Caricaceae. Os resultados desse trabalho sugerem a ocorrência de isoformas de cisteíno-proteases em *Carica candamarcensis*.

O clone E3 possui um sítio de glicosilação na posição ⁴⁴NLT₄₆ sugerindo a semelhança dessa seqüência com a de outra protease de *C. candamarcensis*, CC-III.

O clone escolhido para a expressão, denominado F2, possui 27 substituições de aminoácidos em relação a CMS2MS2, a maioria delas conservativas. Essa seqüência mantém regiões altamente conservadas entre cisteíno-proteases como ¹¹GAVTPV₁₆, ⁷³IKNSWG₁₇₈, os resíduos do sítio ativo C₂₅, H₁₅₉ e N₁₇₅ e as cisteínas formadoras de ponte dissulfeto nas posições 22-63, 56-95 e 153-200.

A sub-clonagem do inserto F2 em plasmídeo pET28a(+)-TEV resultou numa proteína de 292 aminoácidos, 78 resíduos a mais do que a seqüência de CMS2MS2 extraída do látex de *C. candamarcensis*, por conter regiões do vetor.

A expressão do clone F2 no vetor pET28a(+)-TEV, em bactérias BL21-DE3, resultou na expressão de uma proteína com aproximadamente 30 kDa, presente somente na fração insolúvel. A análise *in silico* dos parâmetros bioquímicos da proteína recombinante demonstrou que ela possui pI semelhante ao de CMS2MS2.

A purificação da proteína recombinante em coluna de afinidade ao níquel teve um rendimento de 35 mg de proteína por litro de cultura bacteriana.

A proteína recombinante dissolvida em tampão rico em uréia apresentou atividade mitogênica sobre fibroblastos L929 nas concentrações de 0,4 e 1 nM e em

osteoblastos obtidos de cultura primária na concentração de 1 nM. Os dados sugerem que determinantes lineares, em detrimento dos conformacionais, poderiam ser responsáveis pela atividade mitogênica apresentada por algumas cisteíno-proteases de *Carica candamarcensis*.

7. Referências Bibliográficas

- Antonaccio, M.J.; Normandin, D.; Serafino, R.; Moreland, S. Effects of thrombin receptor activating peptides on rat aortic vascular smooth muscle. *J. Phamacol. Exp. Ther.*, v. 266, p. 125-132, 1993.
- Ausubel, F.; Kingston, R.; Moore, D.; Seidman, J.; Smith, J.; Struhl, K. Unit 4.3. *Current Protocols in Molecular Biology*. John Wiley and Sons Inc., New York, 1987.
- Azarkan, M.; Moussaoui, A.; Wuytswinkel, P.; Dehong, G.; Looze, Y. Fractionation and purification of the enzymes stored in the latex of *Carica papaya*. *J. Chromat. B*, v. 790, p. 229-238, 2003.
- Azarkan, M.; Wintjens, R.; Looze, Y.; Volant, D. B. Detection of three wound-induced proteins in papaya latex. *Phytochem.*, v. 65, p. 525-534, 2004.
- Badillo, V. M. *Carica L. vs. Vasconcellea St. Hil.* (Caricaceae) con la rehabilitación de este último. *Ernstia*, v. 10, p. 74-79, 2000.
- Baeza, G.; Correa, D.; Salas, C.E. Proteolytic enzymes in *Carica candarmacensis*. *J. Sci. Food Agric.*, v. 51, p. 1-9, 1990.
- Bajaj, M.; Waterfield, M. D.; Schlessinger, J.; Taylor, W.R.; Blundell, T. On the tertiary structure of the extracellular domains of the Epidermal Growth Factor and insulin receptors. *Biochem. Biophys. Acta*, v. 916, p. 220-226, 1987.
- Baker K. C.; Taylor M. A.; Cummings N. J.; Tuñón M. A.; Worboys K. A.; Connerton I. F. Autocatalytic processing of pro-papaya proteinase IV is prevented by crowding of the active-site cleft. *Protein Eng.*, v. 9, p. 525-529, 1996.
- Bauvois, B. Transmembrane proteases in cell growth and invasion: new contributors to angiogenesis? *Oncogene*, v. 23, p. 317-329, 2004.
- Bazan, J.F.; Fletterick, R.J. Viral cysteine proteinases are homologous to the trypsin-like family of serine proteases: structural and functional implication. *Proc. Natl. Aca. Sci. USA*, v. 85, p. 7872-7876, 1988
- Biswas, S.; Chakrabarti, C.; Kundu, S.; Jagannadham, M.V.; Dattagupta, J.K. Proposed amino acid sequence and the 1.63 Å X-ray crystal structure of a plant cysteine protease, ervatamin B: Some insights into the structural basis of its stability and substrate specificity. *Proteins: Struct., Funct., Genet.*, v. 51, p. 489-497, 2003.
- Blackman, M. J. Proteases in host cell invasion by the malaria parasite. *Cell. Microbiol.*, v. 6, p. 893-903, 2004.
- Brass, L.F.; Molino, M. Protease-activated G protein-coupled receptors on human platelets and endothelial cells. *Thromb. Haemost.*, v. 78, p. 234-241, 1997.

Bromme, D.; Nallaseth, F.S.; Turk, B. Production and activation of recombinant papain-like cysteine proteases. *Methods*, v. 32, p. 199-206, 2004.

Bryan, P. N. Prodomains and protein folding catalysis. *Chem. Rev.*, v. 102, p. 4805-4815, 2002.

Caffrey, C.R.; Hansell, E.; Lucas, K.D.; Brinen, L.S.; Alvarez Hernandez, A.; Cheng, J.; Gwaltney, S.L.; Roush, W.R.; Stierhof, Y.D.; Bogyo, M.; Steverding, D.; McKerrow, J.H. Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of *Trypanosoma brucei* rhodesiense. *Mol. Biochem. Parasitol.*, v. 118, p. 61-73, 2001.

Carpenter, G.; Cohen, S. Epidermal Growth Factor. *J. Biol. Chem.*, v. 265, p. 7709-7712, 1990.

Chapman, H. A.; Riese, R. J.; Shi, G-P. Emerging roles for cysteine proteases in human biology. *Annu. Rev. Physiol.*, v. 59, p. 63-88, 1997.

Chen, Y.H.; Pouyssegur, J.; Courtneidge, S.A.; van Obberghen-Schilling, E. Activation of SRC family kinase activity by the G protein-coupled thrombin receptor in growth-responsive fibroblasts. *J. Biol. Chem.*, v. 269, p. 27372-27377, 1994.

Choe, Y.; Brinen, L.S.; Price, M.S.; Engel, J.C.; Lange, M.; Grisostomi, C.; Weston, S.G.; Pallai, P.V.; Cheng, H.; Hardy, L.W.; Hartsough, D.S.; McMakin, M.; Tilton, R.F.; Baldino, .M.; Craik, C.S. Development of alpha-keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease. *Bioorg. Med. Chem.*, v. 13, p. 2141-2156, 2005.

Cohen, L. W.; Coghlan, V. M.; Dihel, L. C. Cloning and sequencing of papain-encoding cDNA. *Gene*, v. 48, p. 219-227, 1986.

Cohen, L. W.; Fluharty, C.; Dihel, L. C. Synthesis of papain in *Escherichia coli*. *Gene*, v. 88, p. 263-267, 1990.

DeClerck, Y.A.; Mercurio, A.M.; Stack, M.S.; Chapman, H.A.; Zutter, M.M.; Muschel, R.J.; Raz, A.; Matrisian, L.M.; Sloane, B.F.; Noel, A.; Hendrix, M.J.; Coussens, L.; Padarathsingh, M. Proteases, Extracellular Matrix, and Cancer. *Am. J. Pathol.*, v. 164, p. 1131-1139, 2004.

Denadai, A. M.; Santoro, M. M.; Salas, C. E.; Lopes, M. T.; Avelar, G. M.; Guzman, F. A supramolecular complex between proteinases and beta-cyclodextrin that preserves enzymatic activity: physicochemical characterization. *BioDrugs* (Auckland), v. 20, p. 287-291, 2006.

Déry, O.; Corvera, C.U.; Steinhoff, M.; Bunnett, N.W. Proteinase-activated receptors: novel mechanisms of signaling by serine proteases. *Am. J. Physiol.*, v. 274, p. C1429-C1452, 1998.

Drenth, J., Jansonius, J.N., Koekoek, R., Swen, H.M. & Wolthers, B.G. Structure of papain. *Nature*.v. 218, p. 929-932, 1968.

Duke, J. A.; duCellier, J. L. CRC Handbook of Alternative Cash Crops, Boca Raton, CRC Press, 1993, 533 p.

Eckert, K. A.; Kunkel, T. A. DNA polymerase fidelity and the polymerase chain reaction. *Genome Res.*, v. 1, p. 17-24, 1991.

Gavira, J.A.; González-Ramírez, L.A.; Oliver-Salvador, M.C.; Soriano-García, M.; García-Ruiz, J.M. Structure of the mexicain-E-64 complex and comparison with other cysteine proteases of the papain family. *Acta Crystallogr D Biol Crystallogr.*, v. 63, p. 555-563, 2007.

Gomes, M. T. R. Caracterização bioquímica e farmacológica de proteases do látex de *Carica candamarcensis* com atividade mitogênica. 2008. 146 f. Tese (Doutorado em Bioquímica e Imunologia). Universidade Federal de Minas Gerais, Belo Horizonte.

Gomes, M.T.R.; Bemquerer, M.P.; Lopes, M.T.; Richardson, M.; Júnior, S.O; Salas, C.E. The structure of CMS2MS2, a mitogenic protein isolated from *Carica candamarcensis*. *Biol. Chem.*, v. 388, p. 819-82, 2007.

Gomes, M.T.R.; Mello, V.J.; Rodrigues, K.C.; Bemquerer, M.P.; Lopes, M.T.; Faça, V.M.; Salas, C.E. Isolation of two plant proteinases in latex from *Carica candamarcensis* acting as mitogens for mammalian cells. *Planta Med.*, v. 71, p. 244-248, 2005.

Gomes, M.T.R.; Teixeira, R.D.; Ribeiro, H.A.; Turchetti, A.P.; Junqueira, C.F.; Lopes, M.T.; Salas, C.E.; Nagem, R.A. Purification, crystallization and preliminary X-ray analysis of CMS1MS2: a cysteine proteinase from *Carica candamarcensis* látex. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, v. 64, p. 492-494, 2008.

Gravina-de-Moraes, M.; Termignoni, C.; Salas, C.E. Biochemical characterization of a new cysteine endopeptidase from *Carica candamarcensis* L. *Plant Sci.*, v. 102, p. 11-18, 1994.

Groves, M. R.; Taylor, M. A. J.; Scott, M.; Cummings, N. J.; Pickersgill, R. W.; Jenkins, J. A. The prosequence of procaricain forms an alfa-helical domain that prevents access to the substrate-binding cleft. *Structure*, v. 4, p. 1193-1203, 1996.

Grzonka, Z.; Jankowska, E.; Kasprzykuwski, F.; Harsprzykowska, R.; Lankilwicz, L.; Wicz, W.; Wiczerzak, E.; Ciarkowski, J.; Drabik, P.; Janowski, R.; Kozak, M.; Jaskolski, M.; Grubb, A. Structural studies of cysteine proteases and their inhibitors. *Acta Biochem. Polonica*, v. 48, p. 1-20. 2001.

Harding, A.; Giles, N.; Burgess, A.; Hancock, J.F.; Gabrielli, B.G. Mechanism of mitosis-specific activation of MEK1. *J. Biol. Chem.*, v. 278, p. 16747-16754, 2003.

Hengen, P. N. Purification of His-Tag fusion proteins from *Escherichia coli*. *Trends Biochem. Sci.*, v. 20, p. 285-286, 1995.

Hupp, T.R.; Lane, D.P.; Ball, K.L. Strategies for manipulation the p53 pathway in the treatment of human cancer. *Biochem. J.*, v. 352, p. 1-17, 2000.

Iwasaki, H.; Eguchi, S.; Schichiri, M.; Maruno, F.; Hirata, Y. Adrenomedulin as a novel growth-promoting factor for cultured vascular smooth muscle cells: role of tyrosine kinase-mediated mitogenic-activated protein kinase activation. *Endocrinol.*, v. 138, p. 3432-3441, 1998.

Jaziri M.; Kleinschmidt T.; Walraevens V.; Schnek A.G.; Looze Y. Primary structure of CC-III, the glycosylated cysteine-proteinase from the latex of *Carica candamarcensis* Hook. *Biol Chem Hoppe Seyler*, v. 375, p. 379-385, 1994.

Jia, Z.; Hasnain, S.; Hiramata, T.; Lee, X.; Mort, J.S.; To, R.; Huber, C.P. Crystal structures of recombinant rat cathepsin B and a cathepsin B-inhibitor complex. Implications for structure-based inhibitor design. *J. Biol. Chem.*, v. 270, p. 5527-5533, 1995.

Junqueira, C. F. Caracterização molecular de cisteíno proteases de *Carica candamarcensis*. 2007. Dissertação (Mestrado em Bioquímica e Imunologia). Universidade Federal de Minas Gerais, Belo Horizonte.

Konno, K.; Hirayama, C.; Hirayama, C.; Nakamura, M.; Tateishi, K.; Tamura, Y.; Hattori, M.; Kohno, K. Papain protects papaya trees from herbivorous insects: role of cysteine proteases in latex. *Plant J.*, v. 37, p. 370-378, 2004.

Krausslich, H.C.; Wimmer, E. Viral proteinases. *Ann. Ver. Biochem.*, v. 57, p. 701-754, 1988.

Kyndt, T.; Van Damme, .E.J.; Van Beeumen, J.; Gheysen, G. Purification and characterization of the cysteine proteinases in the latex of *Vasconcellea* spp. *FEBS J.*, v. 274, p. 451-462, 2007.

Kyndt, T.; Van Droogenboerk, B.; Romeijn-Peters, E.; Romero-Motochi J. P.; Scheldeman, X.; Goetghebeur P.; Van Damme, P.; Gheysen, G. Molecular phylogeny and evolution of Caricaceae based on rDNA internal transcribed spacers and chloroplast sequence data. *Mol. Phylogenet. Evol.*, v. 37, p.442-459, 2005.

Laemmli, U.K. Protein in electrophoresis in slab gels. *Nature*, v. 227, p. 680-683, 1970.

Leof, E. B.; Wharton, W.; van Wyh, J. J.; Pledger, W. J. Epidermal growth factor (EGF) and somatomedin C regulate G1 progression in competent BALB/c-3T3. *Exp. Cell Res.*, v. 141, p. 107-115, 1982.

Maes, D.; Bouckaert, J.; Poortmans, F.; Wyns, L.; Looze, Y. Structure of chymopapain at 1.7 Å resolution. *Biochemistry*, v. 35, p. 16292-16298, 1996.

Mahmoudzadeh-Niknam, H.; McKerrow, J.H. Leishmania tropica: cysteine proteases are essential for growth and pathogenicity. *Exp. Parasitol.*, v. 106, p. 158-163, 2004.

Massagué, J.; Pandiella, A. Membrane-anchored growth factors. *Ann. Rev. Biochem.*, v. 62, p. 515-541, 1993.

Mello, V. J. Avaliação da atividade cicatrizante gástrica/cutânea e de parâmetros toxicológicos da fração purificada P1G10 do látex obtido de *Carica candamarcensis*. 2005. Tese (Doutorado em Ciências Biológicas (Fisiologia e Farmacologia)). Universidade Federal de Minas Gerais, Belo Horizonte.

Mello, V.J.; Gomes, M.T.R.; Lemos, F.O.; Delfino, J.L.; Andrade, S.P.; Lopes, M.T.P.; Salas, C.E. The gastric ulcer protective and healing role of cysteine proteinases from *Carica candamarcensis*. *Phytomedicine*, v. 15, p. 237-244, 2008.

Mello, V.J.; Gomes, M.T.R.; Rodrigues, K.C.L.; Sanchez, E.F.; Lopes, M.T.; Salas, C.E. Plant proteinases: their potential as therapeutic drugs, v. 11, p. 213-224. In: *Recent progress in medicinal Plants: Development from new molecules*, Studium Press LLC, Houston, 2006.

Mitchison, T.J. Microtubule dynamics and kinetochore function in mitosis. *Ann. Rev. Cell Biol.*, v. 4, p. 527-549, 1988.

Moses, M.A. The regulation of neovascularization of matrix metalloproteinases and their inhibitors. *Stem Cells*, v. 15, p. 180-189, 1997.

Moussaoui, A.; Nijs, M.; Paul, C.; Wintjens R.; Vincentelli J.; Azarkan, M.; Looze, Y. Revisiting the enzymes stored in the laticifers of *Carica papaya* in the context of their possible participation in the plant defence mechanism. *Cell. Mol. Life Sci.*, v. 58, p. 556-570, 2001.

Moutim, V.; Silva, L.G.; Lopes, M.T.P.; Fernandes G.W.; Salas, C.E. Spontaneous processing of peptides during coagulation of latex from *Carica papaya*. *Plant Sci.*, v. 142, p. 115-121, 1999.

Murray, A.W.; Szostak, J.K. Chromosome segregation in mitosis and meiosis. *Ann. Rev. Cell Biol.*, v. 1, p. 289-315, 1985.

Mynott, T.L.; Ladhams, A.; Scarmato, P.; Engwerda, C.R. Bromelain, from pineapple stems, proteolytically blocks activation of extracellular regulated kinase-2 in T cells. *J. Immunol.*, v. 163, p. 2568-2575, 1999.

Nurse, P. Universal control mechanisms regulation onset of M phase. *Nature*, v. 344, p. 503-508, 1990.

O'Hara, B. P.; Hemmings, A. M.; Buttle, D. J.; Pearl, L. H. Crystal structure of glycyl endopeptidase from *Carica papaya*: a cysteine endopeptidase of unusual substrate specificity. *Biochemistry*, v. 40, p. 13190-13195, 1995.

- Obaru, K. A cDNA clone used to study mRNA inducible in human tonsillar lymphocytes by tumor promoter. *J. Biochem.(Tokyo)*, v. 99, p. 885-94, 1986.
- Ogiso, H.; Ishitani, R.; Nureki, O.; Fukai, S.; Yamanaka, M.; Kim, J.; Saito, K.; Sakamoto, A.; Inoue, M.; Shirouzu, M.; Yokoyama, S. Crystal structure of the complex of the human Epidermal Growth Factor and receptor extracellular domains. *Cell*, v. 110, p. 775-787, 2002.
- Ortega, A.; Pino, J.A. Los constituyentes volátiles de las frutas tropicales. II. Frutas de las especies de *Carica*. Alimentaria. Revista Tecnología Higiene Alimentos, v. 286, p. 27-40, 1997.
- Oshima, J.; Campisi, J. Fundamentals of cell proliferation: control of the cell cycle. *J. Dairy Sci.*, v. 74, p. 2778-2787, 1991.
- Ossovskaya, V.S.; Bunnett, N.W. Protease-activated receptors: contribution to physiology and disease. *Physiol. Rev.*, v. 84, p. 579-621, 2004.
- Pardee, A. B. G1 events and regulation of cell proliferation. *Science*, v. 246, p. 603-608, 1989.
- Pardee, A. B.; Blagosklonny, M. The Restriction Point of the Cell Cycle. *Cell Cycle*, v. 1, p. 103-110, 2002.
- Pereira, M.T.; Lopes, M.T.P.; Meira, W.O.; Salas, C.E. Purification of a cysteine ptoteinase from *Carica candamarcensis* L. and cloning of a genomic putative fragment coding for this enzyme. *Prot. Express. Purificat.*, v. 22, p. 249-257, 2001.
- Pérez, G. Isolation and characterization of a novel lectin from *Doclea lehmanni*. *Phytochem.*, v. 29, p. 1745-1749, 1998.
- Phares, K. An unusual host-parasite relationship: the growth hormone-like factor from plerocercoids of Spirometrid Tapeworms. *Int. J. Parasitol.*, v. 26, p. 575-588, 1996.
- Philchenkov, A. Caspases: Potential targets for regulating cell death. *J. Cell. Mol. Med.*, v. 8, p. 432-444, 2004.
- Pierson, R. W. Jr.; Temin, H. M. The partial purification from calf serum of a fraction with multiplication-stimulating activity for chicken fibroblasts in cell culture and with non-suppressible insulin-like activity. *J. Cell Physiol.*, v. 79, p. 319-330, 1972.
- Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.*, v. 102, p. 4639-4750, 2002.

- Ramjee, M. K.; Petithory, J. R.; McElver, J.; Weber, S. C.; Kirsch, J. F. A novel yeast expression/secretion system for the recombinant plant thiol endoprotease propapain. *Protein Engng.*, v. 9, p. 1055-1061, 1996.
- Rawlings, N.D.; Morton, F.R.; Kok, C.Y.; Kong, J.; Barrett, A.J. *MEROPS*: the peptidase database. *Nucleic Acids Res.*, v. 36, p. D320-D325, 2008.
- Rawlings, N.D.; Barrett, A. J. Evolutionary families of peptidases. *Biochem. J.*, v. 290, p. 205–218, 1993.
- Revell, D.F.; Cummings, N. J.; Baker, K. C.; Collins, M. E.; Taylor, M.A.; Sumner, I. G.; Pickersgill, R. W.; Connerton, I. F.; Goodenough, P.W. Nucleotide sequence and expression in *Escherichia coli* for cDNAs encoding papaya proteinase omega from *Carica papaya*. *Gene*, v. 30, p. 221-225, 1993.
- Robinson, M. W.; Dalton, J. P.; Donnelly, S. Helminth pathogen cathepsin proteases: it's a family affair. *Trends Biochem. Sci. Trends in Biochemical Sciences*, v. 33, p. 601-608, 2008.
- Ross, M.; Ballard, F.J. Regulation of protein metabolism and DNA synthesis by Fibroblast Growth Factor. *Biochem. J.*, v. 249, p. 363-368, 1988.
- Rozengurt, E. Early signals in the mitogenic response. *Science*, v. 234, p. 161-166, 1986.
- Sajid, M.; McKerrow, J.H. Cysteine proteases of parasitic organisms. *Mol. Biochem. Parasitol.*, v. 120, p. 1-21, 2002.
- Salas, C.E.; Gomes, M.T.R.; Hernandez, M.; Lopes, M.T.P. Plant cysteine proteinases: evaluation of the pharmacological activity. *Phytochemistry*, v. 69, p. 2263-2269, 2008.
- Sanger, F.; Nicklen, S.; Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Aca. Sci. USA*, v. 74, p. 5463-5467, 1977.
- Seipelt, J.; Guarné, A.; Bergmann, E.; James, M.; Sommergruber, W.; Fita, I.; Skern, T. The structures of picornalviral proteinases. *Virus Res*, v. 62, p. 159-168, 1999.
- Shindo, T.; Van der Hoorn, R. A. L. Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. *Mol. Plant Pathol.*, v. 9, p. 119-125, 2008.
- Silva, L.G., Garcia, O., Lopes, M.T.P. & Salas, C.E. Changes in protein profile during coagulation of latex of *Carica papaya*. *J. Med. Bio. Res.*, v. 30, p. 615-619, 1997.
- Stiles, C. D.; Capone, G. T.; Schert, C. D.; Antoniades, H. N.; van Wykt, J. J.; Pledgert, W. J. Dual control of cell growth by somatomedins and platelet-derived growth factor. *Cell Biology*, v. 76, p. 1279-1283, 1979.

Storer, A.C.; Menard, R. Catalytic mechanism in papain family of cysteine peptidases. *Methods Enzymol.*, v. 244, p. 486-500, 1994.

Taylor, M. A. J.; Pratt, K. A.; Revell, D.F.; Baker, K. C.; Sumner, I. G.; Goodenough, P. W. Active papain renatured and processed from insoluble recombinant propapain expressed in *Escherichia coli*. *Protein Engng.*, v. 5, p. 455-459, 1992.

Taylor, M. A. J.; Baker, K. C.; Briggs, G. S.; Connerton, I. F.; Cummings, N. J.; Pratt, K. A.; Revell, D.F.; Freedman, R. B.; Goodenough, P. W. Recombinant pro-regions from papain and papaya proteinase IV are selective high affinity inhibitors of the mature papaya enzymes. *Protein Eng.*, v.8, p.59-62, 1995.

Taylor, M. A. J.; Al-Sheikh, M.; Revell, D. F.; Sumner, I. G.; Connerton, I. F. cDNA cloning and expression of *Carica papaya* prochymopapain isoforms in *Escherichia coli*. *Plant Science*, v. 145, p. 41-47, 1999.

Teixeira, R.D.; Ribeiro, H.A.L.; Gomes, M.T.R.; Lopes, M.T.; Salas, C.E. The proteolytic activities in latex from *Carica candamarcensis*. *Plant Physiol. Biochem.*, v. 46, p. 956-961, 2008.

Travis, J.; Potempa, J.; Maeda, H. Are bacterial proteinases pathogenic factors? *Trends Microbiol.*, v. 3, p. 405-407, 1995.

Turk, D.; Guncar, G.; Podobnik, M.; Turk, B. Revised definition of substrate binding sites of papain-like cysteine proteases. *Biol. Chem.*, v. 379, p.137-147, 1998.

Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dult, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Yarden, Y.; Libermann, T.A.; Schlessinger, J. Human Epidermal Growth Factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature*, v. 309, p. 418-425, 1984.

Van der Hoorn, R. A. L. Plant proteases: from phenotypes to molecular mechanisms. *Annu. Rev. Plant Biol.*, v. 59, p. 191-223, 2008.

Van Droogenbroeck, B.; Kyndt, T.; Maertens, I.; Romeijn-Peeters, E.; Scheldeman, X.; Romero-Motochi, J.P.; Van Damme, P.; Goetghebeur, P.; Gheysen, G. Phylogenetic analysis of the highland papayas (*Vasconcellea*) and allied genera (Caricaceae) using PCR-RFLP. *Theor Appl Genet.*, v. 108, p.1473-1486, 2004.

Van Droogenbroeck, B.; Kyndt, T.; Romeijn-Peeters, E.; van Thuyne, W.; Goetghebeur, P.; Romero-Motochi, J.P.; Gheysen, G. Evidence of natural hybridization and introgression between *Vasconcellea* species (Caricaceae) from Southern Ecuador revealed by chloroplast, mitochondrial and nuclear DNA markers. *Annals of Botany*, v. 97, p. 793-805, 2006.

Varughese, K.I.; Su, Y.; Cromwell, D.; Hasnain, S.; Xuong, N.H. Crystal structure of an actinidina-E-64 complex. *Biochemistry*, v. 31, p. 5172-5176, 1992.

Vernet, T.; Khouri, H. E.; Laflamme, P.; Tessier, D. C.; Musil, R.; Gour-Salin, B. J.; Storer, A. C.; Thomar, D. Y. Processing of the papain precursor: purification of the zymogen and characterization of its mechanism of processing. *J. Biol. Chem.*, v. 266, p. 21451-21457, 1991.

Vu, T.K.; Hung, D.T.; Wheaton, V.I.; Coughlin, S.R. Molecular cloning of a novel proteolytic mechanism of receptor activation. *Cell*, v. 64, p. 1057-1068, 1991.

Wahl, M.J. & Carpenter, G. Role of growth factors and their receptors in the control of normal cell proliferation and cancer. *Clin. Physiol. Biochem.*, v. 5, p. 130-139, 1987.

Walraevens, V.; Vandermeers-Piret, M.; Vandermees, P.; Goulert, P. Robberecht, P. Isolation and characterization of the CCI papain-like cysteine proteinases from the latex of *Carica candamarcensis* Hook, *Biol. Chem.*, v. 380, p. 485-488, 1999.

Ward, C.W.; Hoyne, P.A.; Flegg, R.H. Insulin and Epidermal Growth Factor receptors contain the cysteine repeat motif found in the Tumor Necrosis Factor receptor. *Proteins*, v. 22, p. 141-153, 1995.

Weinert, T.; Hartwell, L.H. The *RAD9* gene controls the cycle response to DNA damage in *S. cerevisiae*. *Science*, v. 241, p. 317-322, 1988.

Wiederanders, B. Structure-function relationships in class CA1 cysteine peptidases propeptides. *Acta Biochem. Polonica*, v. 50, p. 691-713, 2003.

Yamamoto, D.; Matsumoto, K.; Ohishi, H.; Ishida, T.; Inoue, M.; Kitamura, K.; Mizuno, H. Refined x-ray structure of papain-E-64-c complex at 2.1 Å resolution. *J. Biol. Chem.*, v. 266, p. 14771-14777, 1991.

Zetter, B.R.; Chen, L.B.; Buchanan, J.M. Effects of protease treatment on growth, morphology, adhesion, and cell surface proteins of secondary chick embryo fibroblasts. *Cell*, v. 7, p. 407-412, 1976.

Zhao, B.; Janson, C.A.; Amegadzie, B.Y.; D'Alessio, K.; Griffin, C.; Hanning, C.R.; Jones, C.; Kurdyla, J.; McQueney, M.; Qiu, X.; Smith, W.W.; Abdel-Meguid, S.S. Crystal structure of human osteoclast cathepsin K complex with E-64. *Nat. Struct. Biol.*, v. 4, p. 109-111, 1997.

