

**Livia José Regatieri**

**Expressão, purificação e caracterização estrutural de proteínas  
codificadas por dois fitovírus**

São José do Rio Preto  
2014

**Livia José Regatieri**

**Expressão, purificação e caracterização estrutural de proteínas  
codificadas por dois fitovírus**

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao Programa de Pós-Graduação em Microbiologia, Área de Concentração – Bioquímica e Biologia Molecular de Microrganismos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

**Orientador:** Prof. Dr. Raghuvir  
Krishnaswamy Arni

São José do Rio Preto  
2014

Regatieri, Livia José.

Expressão, purificação e caracterização estrutural de proteínas codificadas por dois fitovírus / Livia José Regatieri. – São José do Rio Preto, 2014

107 f. : il., gráfs., tabs.

Orientador: Raghuvir Krishnaswamy Ami

Tese (doutorado) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências, Letras e Ciências Exatas

1. Biologia molecular. 2. Vírus de plantas. 3. Carlavíroses. 4. Clonagem molecular. 5. Redobramento de proteína. 6. Luz – Espalhamento. I. Ami, Raghuvir Krishnaswamy. II. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU – 576.858.8

Ficha catalográfica elaborada pela Biblioteca do IBILCE  
UNESP - Campus de São José do Rio Preto

Livia José Regatieri

Expressão, purificação e caracterização estrutural de proteínas  
codificadas por dois fitovírus

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao Programa de Pós-Graduação Microbiologia, Área de Concentração – Bioquímica e Biologia Molecular de Microrganismos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

**Comissão Examinadora**

Prof. Dr. Raghuvir Krishnaswamy Arni  
UNESP – São José do Rio Preto  
Orientador

Profa. Dra. Andréia Estela Moreira de Souza  
FUNEC – Santa Fé do Sul

Profa. Dra. Priscila Belintani  
UNORP – São José do Rio Preto

Prof. Dr. Anwar Ullah  
UNESP – São José do Rio Preto

Prof. Dr. José Ramon Beltran Abrego  
UNESP – São José do Rio Preto

São José do Rio Preto  
03 de fevereiro de 2014

## ***DEDICATÓRIA***

“A todos que direta ou indiretamente contribuíram para a realização deste trabalho.”

## **AGRADECIMENTOS**

Em especial ao meu orientador, Prof. Dr. Raghuvir Krishnaswamy Arni, pela oportunidade de aprendizagem, aprimoramento do conhecimento científico, pela confiança em mim depositada, pelo apoio na continuidade da carreira profissional e, principalmente pela paciência.

Às Prof<sup>as</sup>. Dr<sup>as</sup>. Andréia Estela Moreira de Souza e Priscila Belintani, por toda contribuição científica durante a Iniciação Científica e Mestrado, pela paciência, dedicação, disponibilidade em participar da banca examinadora e pela amizade.

Ao Prof. Dr. José Ramon Beltran Abrego, pelos ensinamentos, pela paciência, pelo incentivo e pela simpatia de sempre.

Ao Dr. Anwar Ullah, por toda contribuição dada a este trabalho, pela ajuda nos ensaios de cristalização das proteínas, pelo comprometimento, dedicação e pela amizade.

Aos Profs. Drs. Luis Henrique Zanini Branco, Monika Aparecida Conrado, Patick Jack Spencer e Uderlei Doniseti Silveira Covizzi, pela disponibilidade em, também, participar da banca examinadora e pelas contribuições diretas e indiretas neste trabalho.

À coordenadora do Programa de Pós-Graduação em Microbiologia Dra. Eleni Gomes, pelo profissionalismo exemplar, pela dedicação e, principalmente, pelas palavras de motivação.

À Me. Lílian Hiromi Tomonari Yamasaki, por toda contribuição dada a este trabalho, principalmente pela ajuda em bioinformática, e pela amizade.

Ao Prof. Dr. Marinônio Cornélio Lopes, por ter me ajudado nas análises de infravermelho e ao Prof. Dr. Marcelo Fossey, por ter me ajudado nos experimentos de dicroísmo circular.

Ao Prof. Dr. José Osmar Gaspar, que me orientou brilhantemente na Iniciação Científica e Mestrado e me incentivou a dar início ao Doutorado.

Aos Professores e funcionários do Instituto de Biociências Letras e Ciências Exatas de São José do Rio Preto pela formação acadêmica e exemplo de profissionalismo e dedicação.

Aos amigos e companheiros de trabalho do Laboratório de Biologia Molecular e do Laboratório de Fitovirologia: Anwar, Carol, Eliana, Hévila, Luana, Mônica, Rehana, Ricardo, Tadaiti e Vinicius, pelo companheirismo, contribuições neste trabalho e pela amizade.

Aos amigos e companheiros de trabalho do Laboratório Municipal de Sorologia: Adriana, Andréa, Erika, Fabiana, Flávia, Liane, Olter, Sandra e Tânia pelo companheirismo.

Aos meus amigos e familiares que me deram muito incentivo e força para concluir este trabalho, sem vocês nada teria sentido.

À FAPESP, CAPES e CNPQ, pelo incentivo à pesquisa.

**Meus sinceros e eternos agradecimentos!**

*“O correr da vida embrulha tudo.  
A vida é assim: esquenta e esfria,  
aperta, daí afrouxa,  
sossega e depois desinquieta.  
O que ela quer da gente é coragem.”*

(Guimarães Rosa)

*“When you lose small mind  
You free your life”*

(Aerials - System of a Down)

## RESUMO

O *Cole latent virus* (CoLV) é um fitovírus cuja principal hospedeira é a couve comum. Pertence ao gênero *Carlavirus*, que é constituído por vírus que infectam diversos grupos de plantas, incluindo aquelas de grande importância econômica no mundo, como a batata, a soja, o amendoim e o milho. O *Pepper ringspot virus* (PepRSV) foi descrito no Brasil no início dos anos 60, sendo caracterizado como integrante do gênero *Tobravirus*, podendo causar danos a culturas como tomate e pimentão. Até o momento, o ciclo de vida desses vírus não é completamente conhecido. Sabe-se que algumas proteínas virais são essenciais na infecção. Dentre estas, a proteína capsidial (CP), a do movimento (MP) e a helicase (HEL) que ainda não possuem estruturas definidas. Desta forma, o trabalho teve como objetivo analisar a estrutura da proteína capsidial do *Cole latent virus* (CP-CoLV); da proteína do movimento do *Pepper ringspot vírus* (MP-PepRSV) e de um fragmento da helicase do *Pepper ringspot vírus* (HEL-PepRSV). Vetores de expressão, contendo as sequências de interesse, foram transformados. Nos testes de indução de expressão a 37°C observou-se a formação de corpos de inclusão para as três proteínas em questão. Após lavagem dos corpos de inclusão, as proteínas purificadas foram reenoveladas e concentradas. Após reenovelamento, a CP-CoLV e a MP-PepRSV foram analisadas por difração de raios X para verificar o seu conteúdo de estrutura secundária. Essa análise confirmou a eficiência do processo de reenovelamento. A CP-CoLV também foi analisada por espalhamento dinâmico de luz, para estimar a distribuição de tamanho das populações de partículas que estão presentes na solução. Os dados do espalhamento dinâmico de luz mostraram que a proteína encontra-se estável e monodispersa. Dessa forma, dentre as proteínas estudadas, a CP-CoLV apresentou melhores condições para os testes de cristalização. Através da expressão a 18°C, foi possível obter a CP-CoLV em seu estado nativo. A proteína foi purificada sob condições não-desnaturantes e analisada por

espectroscopia de infra-vermelho. Por modelagem molecular por homologia e *Threading*, foi possível elaborar modelos tridimensionais da CP-CoLV completa. A predição por ferramentas de bioinformática também mostrou que a estrutura secundária é predominantemente hélice-*coil* e identificou possíveis sítios de modificações pós-traducionais. Os dados sobre a estrutura da MP, HEL e CP deste estudo poderão contribuir para entender o papel destas no ciclo de infecção dos carlavírus e tobnavírus, para elaborar mecanismos de controle e contribuir para o uso biotecnológico destes vírus.

Palavras-chave: clonagem; carlavírus; tobnavírus; expressão; reenovelamento; dicroísmo circular; espalhamento dinâmico de luz; modelagem molecular.

## **ABSTRACT**

*The Cole latent virus is a plant virus which infects mainly the common cole. It is a member of Carlavirus genus, which contains viruses that infect several plant groups, including those of great economic importance worldwide, as potato, soy, peanut and corn. The Pepper ringspot virus was first described during the 60's in Brazil. It is characterized as member of Tobravirus genus and it can damage to cultures as tomato and pepper. Until now, the viral life cycle is not completely elucidated. It is known that some viral proteins are essential for infection, including the capsid protein (CP), movement protein (MP) and helicase (HEL). Those proteins do not have the structure determined, yet. Therefore, this work aimed the structural study of the capsid protein of Cole latent virus (CP-CoLV); movement protein of Pepper ringspot virus (MP-PepRSV) and a helicase fragment of Pepper ringspot virus (HEL-PepRSV). Expression vectors containing sequences of interest were transformed Tests for protein expression at 37°C resulted in formation of inclusion bodies for the three tested proteins. The proteins present in inclusion bodies were solubilized and purified under denaturing conditions. The purified proteins were refolding by dialysis and concentrated. After refolding CoLV-CP and MP-PepRSV were analyzed by circular dichroism to check its secondary structure content. This analysis confirmed the refolding efficiency. The CP-CoLV was also analyzed by Dynamic Light Scattering to estimate the size distribution of particle populations which are present in the solution. The data of the dynamic light scattering showed that the protein is stable and monodisperse. Thus, among the proteins studied, the CP-CoLV showed better conditions for crystallization trials. When expressed at 18°C, it was possible to obtain the CP-CoLV in its native state. The protein was purified under non-denaturing and analyzed by infra-red spectroscopy. Using homology molecular modeling and Threading, it was the tridimensional models for the complete CP-CoLV were obtained. Predictions by bioinformatics tools indicated that secondary structure is mostly helix-coil and*

*possible sites for post-translational modifications. These structural data on MP, HEL and CP may be useful for understanding the role of proteins in the infection cycle of carlaviruses and tobnaviruses to elaborate control mechanisms and biotechnological usage for these viruses.*

*Keywords: cloning; carlavirus; tobnavirus; expression; refolding; circular dichroism; dynamic light scattering; molecular modeling.*

## Lista de Figuras

<b>Figura 1.</b> Organização do genoma dos carlavírus e possíveis proteínas codificadas .....	25
<b>Figura 2.</b> Micrografia eletrônica mostrando partículas virais alongadas do <i>Cole latent virus</i> .....	25
<b>Figura 3.</b> Organização genômica do RNA 1 dos tobnavírus. ....	29
<b>Figura 4.</b> Organização genômica do RNA 2 dos tobnavírus. ....	29
<b>Figura 5.</b> Micrografia eletrônica de transmissão mostrando partículas longas (L) e curtas (C) do PepRSV.....	30
<b>Figura 6.</b> Micrografia eletrônica de transmissão mostrando corte de célula esporogênica de antera de tomateiro apresentando partículas longas (L) e curtas (C) do PepRSV associadas com mitocôndrias.....	32
<b>Figura 7.</b> A) Movimento célula a célula do TMV e B) do TSWV.....	40
<b>Figura 8.</b> Formação de túbulos em protoplastos infectados com TMV.....	40
<b>Figura 9.</b> Gel nativo de agarose 1% evidenciando fragmento de 920 pb amplificado por PCR referente ao gene da CP do CoLV .....	63
<b>Figura 10.</b> Gel nativo de agarose 1% evidenciando fragmento de 800 pb amplificado por PCR referente ao gene da MP do PepRSV .....	64
<b>Figura 11.</b> Gel nativo de agarose 1% evidenciando fragmento de 845 pb amplificado por PCR referente ao gene da helicase do PepRSV .....	64
<b>Figura 12.</b> Gel nativo de agarose 1% evidenciando digestão enzimática dos vetores de expressão com liberação dos fragmentos de interesse.....	65
<b>Figura 13. (A)</b> SDS-PAGE 12% mostrando o teste de expressão a 37°C para a proteína capsial do CoLV; <b>(B)</b> para a proteína do movimento do PepRSV; <b>(C)</b> para a helicase do PepRSV .....	67
<b>Figura 14.</b> SDS-PAGE 12% mostrando os resultados dos testes de expressão a 18°C e purificação da CP do CoLV em condições nativas.....	68
<b>Figura 15.</b> SDS-PAGE 12% mostrando o resultado da purificação da CP do CoLV em condições desnaturantes. ....	69
<b>Figura 16.</b> SDS-PAGE 12% mostrando os resultados do teste de expressão a 18°C e purificação da MP do PepRSV. ....	69
<b>Figura 17.</b> SDS-PAGE 12% mostrando os resultados do teste de expressão a 18°C e purificação da Helicase do PepRSV. ....	70

<b>Figura 18.</b> Western blot de um gel idêntico ao da Figura 18 A evidenciando a expressão das proteínas de interesse. ....	71
<b>Figura 19.</b> Western blot de um gel idêntico ao da Figura 19 A evidenciando a expressão da proteína de interesse (MP-PepRSV). ....	71
<b>Figura 20.</b> Western blot de um gel idêntico ao da Figura 20 A evidenciando a expressão da proteína de interesse (Hel-PepRSV) .....	72
<b>Figura 21:</b> Gel de poliácridamida (SDS-PAGE) 12% mostrando a expressão em <i>E. coli</i> e purificação em resina de níquel da proteína capsial do CoLV fusionada a cauda de histidina. ....	73
<b>Figura 22.</b> Gráfico da segunda derivada da proteína capsial do CoLV quando purificada de forma nativa.....	74
<b>Figura 23.</b> Gráfico da segunda derivada da proteína capsial do CoLV quando purificada de forma desnaturante com posterior reenovelamento. ....	75
<b>Figura 24.</b> Gráfico da deconvolução espectral da banda de Amida I mostrando as contribuições indicadas pelo gráfico de segunda derivada da proteína do CoLV .....	75
<b>Figura 25.</b> Gráfico da deconvolução espectral da banda de Amida I mostrando as contribuições indicadas pelo gráfico de segunda derivada da proteína do CoLV .....	75
<b>Figura 26.</b> SDS-PAGE 12% mostrando as tentativas de solubilização dos corpos de inclusão (CP CoLV) obtidos a 37°C, com diferentes agentes não desnaturantes. ....	77
<b>Figura 27.</b> SDS-PAGE 12% mostrando as tentativas de solubilização na presença de arginina, dos corpos de inclusão (CP CoLV) obtidos a 18°C. ....	78
<b>Figura 28.</b> SDS-PAGE 12% mostrando as tentativas de solubilização dos corpos de inclusão (MP do PepRSV) obtidos a 37°C, com diferentes agentes não desnaturantes.....	78
<b>Figura 29.</b> SDS-PAGE 12% mostrando as tentativas de solubilização dos corpos de inclusão (HEL do PepRSV) obtidos a 37°C, com diferentes agentes não desnaturantes.....	79
<b>Figura 30.</b> SDS-PAGE 12% mostrando as etapas de lavagem do corpo de inclusão e as etapas de purificação da CP do CoLV .....	80
<b>Figura 31.</b> SDS-PAGE 12% mostrando as etapas de lavagem do corpo de inclusão e as etapas de purificação da MP do PepRSV.....	81
<b>Figura 32.</b> SDS-PAGE 12% mostrando as etapas de lavagem do corpo de inclusão e as etapas de purificação da HEL do PepRSV. ....	81
<b>Figura 33.</b> Análise por dicroísmo circular das proteínas CP-CoLV e MP-PepRSV .....	83
<b>Figura 34.</b> SDS-PAGE 12% mostrando o resultado do processo de concentração das proteínas reenoveladas. ....	84

<b>Figura 35.</b> Fotomicrografia do cristal irregular, resultante do ensaio de cristalização da CP-CoLV .....	87
<b>Figura 36.</b> Gráfico de Ramachandran do modelo gerado dos resíduos 99-265 (A), 1-98 (B) e 266-305(C).....	89
<b>Figura 37.</b> Topologia geral do modelo, mostrando a cadeia principal e a superfície .....	90

## Lista de Tabelas

<b>Tabela 1.</b> Classificação dos fitovírus com base na necessidade da proteína capsial para o movimento .....	38
<b>Tabela 2.</b> Sequências dos oligonucleotídeos utilizados para amplificar os genes em estudo. ....	53
<b>Tabela 3.</b> Cálculo das áreas e respectivas porcentagens das bandas gaussianas obtidas por deconvolução espectral da proteína capsial do CoLV purificada em condições nativas. .....	76
<b>Tabela 4.</b> Cálculo das áreas e respectivas porcentagens das bandas gaussianas obtidas por deconvolução espectral da proteína capsial do CoLV purificada em condiçõesdesnaturantes com posterior reenovelamento. .....	76
<b>Tabela 5.</b> Soluções utilizadas no teste de solubilidade e análise da presença de precipitado. .....	86
<b>Tabela 6.</b> Estruturas proteicas utilizadas como molde para a estrutura tridimensional do CP-CoLV. .....	88
<b>Tabela 7.</b> Resumo do gráfico de Ramachadran em porcentagem dos residuos.....	89

## Lista de Vírus Citados no Texto

AcLV:	<i>Aconitum latent virus</i>	( <i>Carlavirus</i> )
AMV:	<i>Alfafa mosaic virus</i>	( <i>Alfamovirus</i> )
BBScV:	<i>Blueberry scorch virus</i>	( <i>Carlavirus</i> )
CaMV:	<i>Cauliflower mosaic virus</i>	( <i>Carlavirus</i> )
CCMV:	<i>Cowpea chlorotic mottle virus</i>	( <i>Bromovirus</i> )
CLV:	<i>Carnation latent virus</i>	( <i>Carlavirus</i> )
CMV:	<i>Cucumber mosaic virus</i>	( <i>Cucumovirus</i> )
CoLV:	<i>Cole latent virus</i>	( <i>Carlavirus</i> )
CPMV:	<i>Cowpea mosaic virus</i>	( <i>Comovirus</i> )
CpMMV:	<i>Cowpea mild mottle virus</i>	( <i>Carlavirus</i> )
CrVB:	<i>Chrysanthemum virus B</i>	( <i>Carlavirus</i> )
CV2:	<i>Cactus virus 2</i>	( <i>Carlavirus</i> )
DENV:	<i>Dengue virus</i>	( <i>Arbovirus</i> )
GV1:	<i>Garlic virus 1</i>	( <i>Carlavirus</i> )
HBV:	<i>Hepatitis B virus</i>	( <i>Orthohepadnavirus</i> )
HCV:	<i>Hepatitis C virus</i>	( <i>Hepacivirus</i> )
HelVS:	<i>Helenium virus S</i>	( <i>Carlavirus</i> )
HopMV:	<i>Hop mosaic virus</i>	( <i>Carlavirus</i> )
HPV:	<i>Human papiloma virus</i>	( <i>Papilomavirus</i> )
LSV:	<i>Lily symptomless virus</i>	( <i>Carlavirus</i> )
PapMV:	<i>Papaya mosaic virus</i>	( <i>Potexvirus</i> )
PEBV:	<i>Pea early-browning virus</i>	( <i>Tobravirus</i> )
PepRS:	<i>Pepper ringspot virus</i>	( <i>Tobravirus</i> )

PopMV:	<i>Poplar mosaic virus</i>	( <i>Carlavirus</i> )
PPV:	<i>Plum pox virus</i>	( <i>Potyvirus</i> )
PVM:	<i>Potato virus M</i>	( <i>Carlavirus</i> )
PVS:	<i>Potato virus S</i>	( <i>Carlavirus</i> )
RCVMV:	<i>Red clover vein mosaic virus</i>	( <i>Carlavirus</i> )
TEV:	<i>Tobacco etch virus</i>	( <i>Potyvirus</i> )
TMV:	<i>Tobacco mosaic virus</i>	( <i>Tobamovirus</i> )
TRV:	<i>Tobacco rattle virus</i>	( <i>Tobravirus</i> )
TSWV:	<i>Tomato spotted wilt virus</i>	( <i>Tospovirus</i> )
TuMV:	<i>Turnip mosaic virus</i>	( <i>Carlavirus</i> )
VAP:	Vírus do Anel do Pimentão	( <i>Tobravirus</i> )

## Lista de Abreviações e Símbolos

- ADA: ácido iminodiácético
- $\alpha$ : alfa
- $\beta$ : beta
- BCIP/NBT: 5-bromo-4-cloro-3-indolil fosfato/nitro-blue tetrazolio
- BSA: albumina de soro bovino
- $^{\circ}\text{C}$ : graus Celsius
- CAPS: ácido N-ciclohexil-3-aminopropanosulfônico
- CD: dicroísmo circular
- cDNA: DNA complementar
- CP: proteína capsidial
- CRVs: complexos de replicação viral
- Da: Dalton
- DLS: *Dynamic Light Scattering* ou Espalhamento dinâmico de luz
- DNA: ácido desoxirribonucleico
- DO: densidade óptica
- DTT: ditioneitol
- *E. coli*: *Escherichia coli*
- EPPS - ácido N-(2-hidroxietil) piperazina-N-3-propanosulfônico
- FT-IR: *Fourier transform infrared spectroscopy*
- HEL: helicase
- IgG: imunoglobulina G
- IPTG: isopropil- $\beta$ -D-galactosidase
- Kb: quilobases
- KCl: cloreto de potássio
- kDa: quilodalton
- LB: meio Luria-Bertani
- M: molar
- MES: ácido 2-morfolinoetanosulfônico monohidratado
- $\mu\text{g}$ : microgramas
- $\mu\text{m}$ : micrômetro
- $\text{MgCl}_2$ : cloreto de magnésio
- min: minutos
- MOPS: ácido morfolinopropanesulfônico
- MP: proteína do movimento
- MW: *molecular weight* ou peso molecular (Daltons)
- NaCl: cloreto de sódio
- $\text{NaHPO}_4$ : fosfato de sódio
- ng: nanograma

- Ni: níquel
- nm: nanômetro
- nt: nucleotídeos
- NTPs: nucleotídeos trifosfatos
- ORF: *Open Reading Frames* ou fases abertas de leitura
- pb: pares de bases
- PBS: tampão fosfato-salino
- PBS-T: PBS contendo Tween 20
- PCR: *Polymerase chain reaction* ou reação em cadeia da polimerase
- PDB: *Protein Data Bank*
- PEG: polietilenoglicol
- pH: potencial hidrogeniônico
- RdRp: RNA polimerase dependente de RNA
- Rep: replicase
- RFs: formas replicativas
- RNA: ácido ribonucleico
- RPM: rotações por minuto
- S: Svedberg
- SDS: lauril sulfato de sódio ou dodecil sulfato de sódio
- TGB: bloco triplo de genes
- Tris-HCl: tampão tris-base com pH ajustado com HCl
- V: volts
- VNPs: complexos ribonucleoproteicos

# SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	22
1. 1 O Gênero <i>Carlavirus</i> .....	22
1. 1. 1 <i>Cole latent virus</i> (CoLV) .....	24
1. 2 O Gênero <i>Tobravirus</i> .....	26
1. 2. 1 <i>Pepper ringspot virus</i> (PepRSV) .....	30
1. 3 Síntese das proteínas virais durante a infecção .....	32
1. 3. 1 Proteína capsidial – Uma proteína multifuncional .....	34
1. 3. 2 O movimento dos vírus nas plantas e a proteína do movimento.....	37
1. 3. 3. Helicases.....	41
1. 4 Expressão de proteínas heterólogas em bactérias .....	43
1. 5 Ferramentas de bioinformática aplicadas ao estudo de estrutura proteica .....	47
<b>2 JUSTIFICATIVA</b> .....	49
<b>3 OBJETIVOS</b> .....	51
<b>4 MATERIAL E MÉTODOS</b> .....	52
4. 1 Obtenção dos plamídeos recombinantes .....	53
4. 2 Transformação em células competentes.....	53
4. 3 Confirmação da transformação por PCR.....	54
4. 4 Confirmação da transformação por restrição enzimática .....	54
4. 5 Testes de expressão e solubilidade em diferentes temperaturas.....	54
4. 6 Análise em gel de poliacrilamida (SDS-PAGE) .....	55
4. 7 Análise da expressão por western-blot com anti-polihistidina .....	55
4. 8 Purificação das proteínas em condições nativas.....	56
4. 9 Análise da CP do CoLV por espectroscopia de infravermelho (IR) .....	56
4. 10 Diálise das frações eluídas para remoção do imidazol.....	57
4. 11 Testes de solubilização com agentes não-desnaturantes .....	57
4. 12 Purificação das proteínas em condições desnaturantes .....	57
4. 13 Purificação das proteínas em condições desnaturantes e reenovelamento por gradiente de uréia .....	58

4. 14	Reenovelamento das proteínas por diluição .....	58
4. 15	Reenovelamento das proteínas por diálise .....	58
4. 16	Concentração e quantificação das amostras purificadas .....	59
4. 17	Análise das proteínas por dicroísmo circular (CD) .....	59
4. 18	Espalhamento dinâmico de luz (DLS) .....	60
4. 19	Teste de solubilidade da CP do CoLV em diferentes condições .....	60
4. 20	Teste de cristalização das proteínas purificadas .....	61
4. 21	Gel filtração das frações eluídas para remoção do imidazol .....	61
4. 22	Predição de estrutura por ferramentas de bioinformática.....	62
<b>5</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>63</b>
5. 1	Transformação dos plasmídeos recombinantes em células competentes .....	63
5. 2	Testes de expressão e solubilidade .....	65
5. 3	Confirmação da purificação das proteínas de interesse.....	70
5. 4	Reenovelamento das proteínas por gradiente 8-0 M uréia.....	72
5. 5	Espectroscopia de infravermelho (FT-IR) da CP do CoLV purificada em condições nativas e em condições desnaturantes com posterior reenovelamento .....	73
5. 6	Testes de solubilidade com agentes não-desnaturantes .....	76
5. 7	Redução de proteínas não específicas através da lavagem dos corpos de inclusão .....	79
5. 8	Reenovelamento das proteínas e análise por dicroísmo circular .....	82
5. 9	Concentração das proteínas e análise por DLS .....	83
5. 10	Ensaio de cristalização .....	86
5. 11	Análise estrutural da CP-CoLV pelo PredictProtein .....	87
<b>6</b>	<b>CONCLUSÕES .....</b>	<b>91</b>
<b>ANEXO 1</b>	<b>.....</b>	<b>92</b>
<b>ANEXO 2</b>	<b>.....</b>	<b>93</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>.....</b>	<b>94</b>

## 1. Introdução

### 1.1 O Gênero *Carlavirus*

A ocorrência de uma doença causada por um carlavírus foi documentada pela primeira vez por SCHULTZ e FOLSOM (1923) quando verificaram sintomas constituídos por mosaico e enrolamento das folhas em batata. Mais tarde, estes sintomas foram atribuídos ao *Potato virus M* (PVM) (BAGNALL et al., 1959).

Os carlavírus são transmitidos mecanicamente e alguns de maneira não persistente por afídeos (HAGEDORN e HANSON, 1951; ADAMS e BARBARA, 1980; KOENIG et al., 1983). O afídeo *Myzus persicae* é um dos principais vetores para diversos vírus do gênero, tais como o *Carnation latent virus* (CLV) (WETTER e MILNE, 1981), *Chrysanthemum virus B* (CrVB) (HOLLINGS, 1972) e *Hop mosaic virus* (HopMV) (CROWLE et al., 2006). Embora sementes infectadas representem uma importante forma de disseminação, são poucos os exemplos de carlavírus em que a transmissão por esta via possua impacto epidemiológico, sendo relatada para o *Hop mosaic virus*, *Red clover vein mosaic virus* (RCVMV) e para alguns isolados do *Cowpea mild mottle virus* (CpMMV) (ADAMS et al., 2004; XU et al., 2010).

Em geral, o gênero apresenta espécies que induzem poucos sintomas em suas plantas hospedeiras ou permanecem latentes, não causando grandes perdas de ordem econômica. Por esse motivo muitos dos nomes dados às espécies de carlavírus incluem o termo “latente” (FOSTER, 1991). Há, no entanto, exceções, onde os sintomas tornam-se mais severos como no caso de infecção por CpMMV em soja e amendoim (BRUNT e KENTEN, 1973; ALMEIDA et al., 2005; AROGUNDADE et al., 2009) e por RCVMV em ervilha, grão-de-bico e lentilha (LARSEN e WYATT, 2009).

Inicialmente, os carlavírus constituíam um grupo denominado “grupo do Vírus S da batata” (BRANDES e BERCKS, 1965) e posteriormente “grupo do Vírus latente do cravo” – *Carlavirus* – das letras iniciais “**C**arnation **l**atent **v**irus” (HARRISON et al., 1971; MATTHEWS, 1991), sendo essa a

espécie tipo do gênero. O gênero *Carlavirus* pertence à família *Betaflexiviridae* e contém, atualmente, 52 espécies descritas (ICTV, 2012).

As partículas alongadas e flexuosas, medem 10-15 nm de diâmetro e 610-700 nm de comprimento (WETER e MILNE, 1981; IWAKI et al., 1982). A partir de dados para a massa molecular da proteína e conteúdo de ácido nucleico, foi estimado o peso molecular da partícula em  $50-60 \times 10^6$  Da, o que indica que cada partícula contém cerca de 1600 a 1800 subunidades proteicas. Estas partículas possuem coeficiente de sedimentação entre 157 S e 172 S, sendo que seu coeficiente de extinção varia de 2,1 a 2,3 (WETTER e MILNE, 1981). A massa molecular da proteína capsidial, estimada por eletroforese em gel desnaturante de SDS-poliacrilamida, varia de 31 kDa a 39 kDa (BRUNT e KENTEN, 1973; GASPAR e COSTA, 1993; VAN REGENMORTEL, 2000; COHEN et al., 2003).

Determinadas moléculas possuem, em sua estrutura, regiões capazes de induzir a formação de anticorpos, tais regiões são denominadas como determinantes antigênicos ou epitopos (ALMEIDA, 2005). As partículas virais dos carlavírus são altamente imunogênicas, enquanto que dos demais gêneros da família *Betaflexiviridae* são considerados como medianos a fracos imonógenos (KING et al., 2011). Em um estudo das relações antigênicas entre as subunidades da capa proteica e dos vírions intactos de *Lily symptomless virus* (LSV), concluiu-se que a proteína capsidial após renaturação conservava a especificidade e antigenicidade da partícula em sua forma nativa (WANG et al., 2010). Através do software ANTHEPROT, utilizando o método descrito por PARKER et al. (1986) e WELLING et al. (1985), verificou-se que as regiões antigênicas concentravam-se nas regiões aminoterminal e carboxiterminal (WANG et al., 2010).

Os vírus pertencentes ao gênero *Carlavirus* caracterizam-se por possuírem genoma de RNA monopartido, de fita única e positiva, com peso molecular variando de  $2,3$  a  $3,0 \times 10^6$  Da (6,4 - 8,6Kb), sendo que o conteúdo de RNA varia de 5% a 8,5%, entre espécies desse gênero (VAN REGENMORTEL, 2000).

A sequência parcial de nucleotídeos da extremidade 3' do RNA genômico foi determinada para várias espécies do gênero, tais como *Potato virus M*, *Potato virus S* (PVS), *Lily symptomless virus*, *Helenium virus S* (HeIVS)

*Cactus virus 2* (CV2), *Poplar mosaic virus* (PopMV), *Garlic virus 1* (GV1) e *Cowpea mild mottle virus*. A sequência completa dos nucleotídeos foi determinada para algumas espécies como o PVM (ZAVRIEV et al., 1991), *Blueberry scorch virus* (BBScV) (CAVILEER et al., 1994) e *Aconitum latent virus* (AcLV) (FUJI et al., 2002) que possuem, respectivamente, 8534, 8514 e 8657 nucleotídeos, excluindo a sequência poli(A).

O conhecimento sobre a sequência de nucleotídeos e conseqüentemente, sobre a estrutura do genoma, possibilita uma melhor compreensão e inferência acerca das estratégias utilizadas na replicação viral. Assim, a análise das sequências completas de diferentes genomas do grupo permitiu identificar seis fases abertas de leitura (Open Reading Frames- ORFs) (Figura 1) (CAVILEER, 1998; FUJI, et al., 2002).

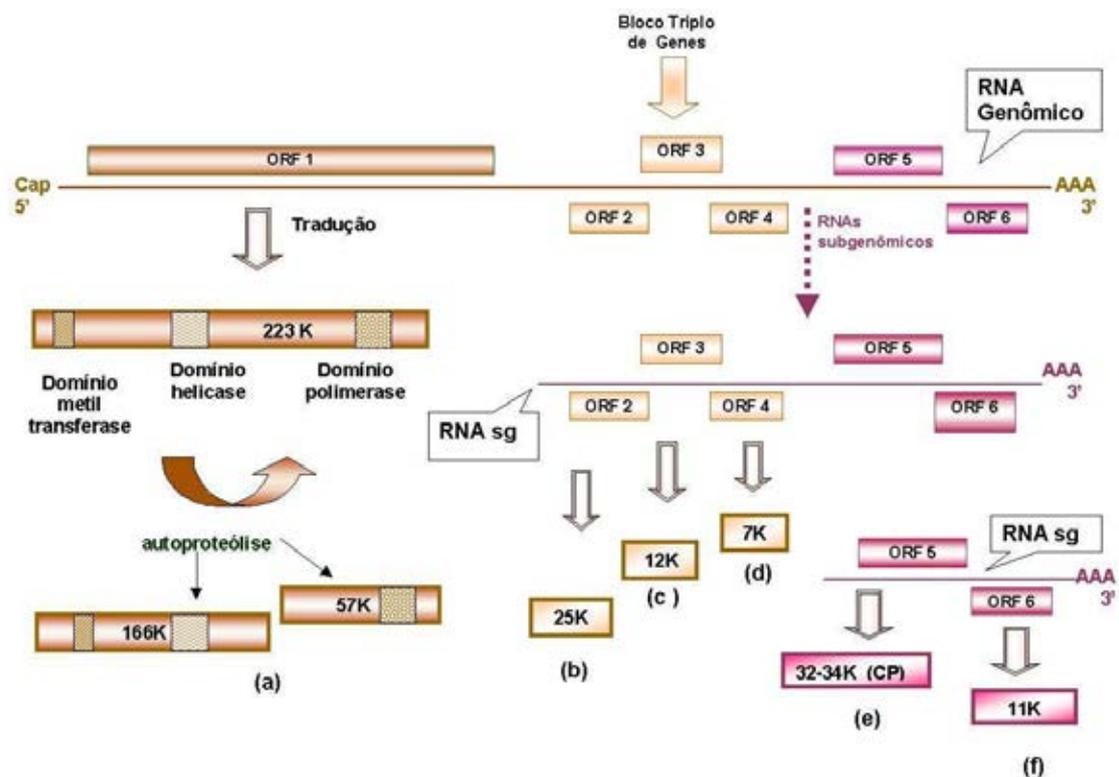
A ORF1 codifica um polipeptídeo com função de metil-transferase, protease, helicase e RNA polimerase dependente de RNA. As ORFs 2, 3 e 4 formam um bloco triplo de genes (TGB) relacionado ao movimento do vírus célula a célula. A ORF5 codifica a proteína capsidial e a região seguinte (ORF 6) codifica uma proteína rica em cisteína, provavelmente, relacionada com a regulação da transcrição viral (EASTWELL e DRUFFEL, 2012; LAWRENCE, et al., 1995).

O RNA genômico possui um 'cap' na extremidade 5' e é poliadenilado na extremidade 3'. Para a expressão dos genes da extremidade 3', os carlavírus utilizam ao menos dois RNAs subgenômicos de 2,1-3,3 Kb e 1,3-1,6 Kb (FOSTER e MILLS, 1991; VAN REGENMORTEL e MAHY, 2009).

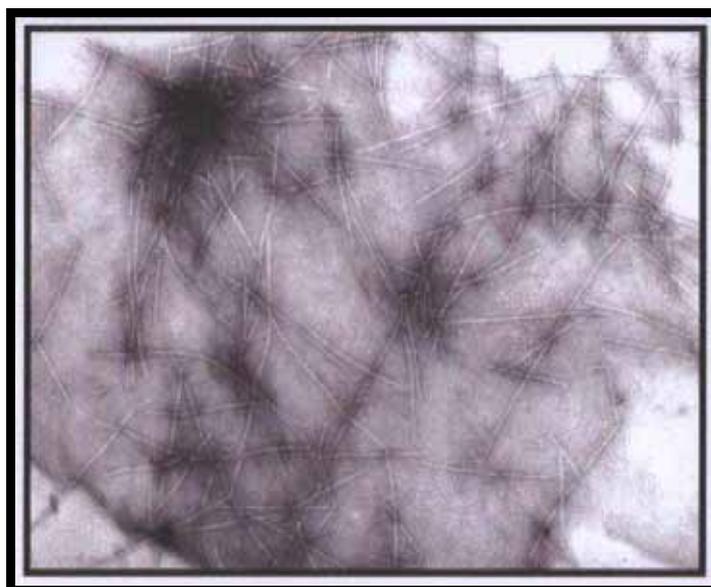
### **1.1.1 *Cole latent virus* (CoLV)**

O *Cole latent virus* (CoLV) é um fitovírus do gênero *Carlavirus*, que apresenta partículas filamentosas (Figura 2) com cerca de 650 nm de comprimento e RNA genômico de 8,3 Kb envolto por moléculas de proteína capsidial com cerca de 34 kDa.

*Cole latent virus* (CoLV) foi descrito pela primeira vez em couve (*Brassica oleracea* L. var. *acephala*) no Brasil, Estado de São Paulo durante pesquisas sobre a suscetibilidade de variedades de couve ao *Turnip mosaic virus* (TuMV) realizadas por COSTA (1961) (dados não publicados).



**Figura 1.** Organização do genoma dos carlavírus e possíveis proteínas codificadas conforme dados obtidos para os vírus BBScV (LAWRENCE et al., 1995) - (a); PVM (GRAMSTAT et al., 1994; RUPASOV et al., 1989; MOROZOV et al., 1991) - (c) (d) e (f), PVS (MACKENZIE et al., 1989) - (b) e (e) HeIVS (FOSTER e MILLS, 1990). As massas moleculares para as proteínas podem variar conforme as espécies do gênero. (Figura obtida de ALMEIDA et al., 2005).



**Figura 2.** Micrografia eletrônica mostrando partículas virais alongadas do *Cole latent vírus* (Fonte: BELINTANI, 2001).

Sua presença foi verificada em praticamente todas as variedades de couve comum ou manteiga estudadas formando um complexo com TuMV e *Cauliflower mosaic virus* (CaMV) ou ocorrendo em infecção isolada (COSTA et al., 1972).

O vírus é transmitido mecanicamente e de maneira não persistente por *Myzus persicae* (MELLO, 1981). Além disso, pode-se averiguar que o CoLV é transmissível por inoculação mecânica e enxertia a várias plantas-teste (A.S. Costa, dados não publicados).

Quanto ao círculo de hospedeiras, MELLO (1981) relata que após a inoculação mecânica, o CoLV infectou todas as crucíferas testadas e uma ou mais espécies das famílias *Amaranthaceae*, *Chenopodiaceae*, *Malvaceae*, *Solanaceae*, *Cucurbitaceae* e *Leguminosae*, sendo que a maioria não desenvolveu sintomas.

KITAJIMA et al. (1970), observando cortes de folhas de couve sistemicamente infectadas, relataram a localização das partículas do CoLV no citoplasma não se achando especificamente associadas a nenhuma organela celular, embora, com certa frequência, as partículas fossem encontradas nas proximidades dos cloroplastos.

Evidências moleculares foram obtidas por BELINTANI et al. (2002) contribuindo para o reconhecimento do CoLV como uma espécie distinta dentro do gênero *Carlavirus*. O RNA genômico do CoLV é poliadenilado mas não foi detectado RNA subgenômico no RNA total extraído de partículas do CoLV. No entanto, dois RNAs subgenômicos (2,6 e 1,3 Kb) e três RNAs de fita dupla (8,3, 2,6 e 1,3 Kb) são produzidos em plantas infectadas pelo CoLV (FOSTER e MILLS, 1991).

## **1. 2 O gênero *Tobravirus***

O gênero *Tobravirus* é composto por três espécies de fitovírus, *Tobacco rattle virus* (TRV), *Pea early-browning virus* (PEBV) e *Pepper ringspot virus* (PepRSV) (HARRISON, 1973; HARRISON e ROBINSON, 1986; ROBINSON e HARRISON, 1989; ICTV, 2012).

O TRV é considerado a espécie-tipo do gênero, tem ampla distribuição mundial e pode infectar um grande número de espécies de plantas causando prejuízos, principalmente, às culturas de batata, fumo e plantas ornamentais. O PEBV foi encontrado no Norte da Europa e Norte da África, onde infecta, primariamente, leguminosas (MACFARLANE, 1999). Já o PepRSV foi descrito somente no Brasil causando doenças em pimentão e tomate (COSTA et al., 1961).

Os vírus do gênero *Tobravirus* apresentam o genoma dividido em dois RNAs de fita única com polaridade positiva, sendo que cada fita de RNA é encapsidada separadamente em partículas tipo bastonetes com dois comprimentos predominantes. As partículas longas possuem comprimento variável de 185 a 212 nm e contêm o RNA 1, enquanto as partículas curtas possuem um comprimento maior, variando de 50 a 115 nm e contêm o RNA 2 (HARRISON e ROBINSON, 1986).

As moléculas de RNA 1 (Figura 3) dos três vírus do gênero possuem tamanhos similares: isolado SYM do TRV, 6791 nucleotídeos (nt); PepRSV, 6828 nt; isolado SP5 do PEBV, 7073 nt. A organização do genoma é conservada, mas a similaridade entre as sequências das três espécies é de apenas 58-65%. O RNA 1 codifica quatro proteínas não estruturais: I) Proteína de 134-141 kDa (domínio Helicase, 'Hel'); II) Proteína de 194-201 kDa (domínios Helicase e RNA polimerase dependente de RNA, 'RdRp') produzida por leitura através do códon de terminação. O gene completo Hel/RdRp, também denominado Replicase, ocupa aproximadamente 75% do RNA 1 e, está relacionado com a replicação do RNA viral; III) Proteína de 29-30 kDa (provável proteína do movimento 'MP' ou Proteína 1a) que mostra alguma homologia com a proteína de 30 kDa codificada pelo *Tobacco mosaic virus* (TMV), a qual é, comprovadamente, responsável pelo movimento célula a célula desse vírus; IV) Proteína de 12-16 kDa, conhecida como Proteína 1b, provavelmente relacionada com a transmissão dos vírus pelas sementes (MACFARLANE, 1999).

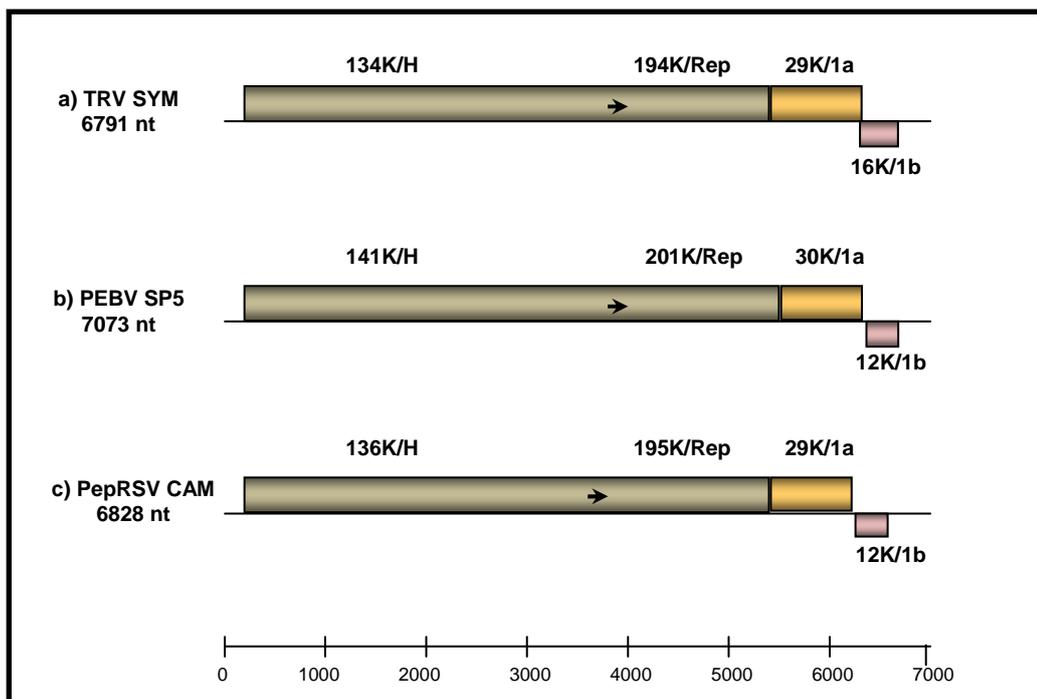
O RNA 2 (Figura 4) é menor que o RNA 1 e varia de 1800 a 3900nt entre os diferentes isolados (BERGH et al, 1985; CORNELISSEN et al., 1986). Apresenta na extremidade 5' um gene que codifica para a proteína capsidial (CP) de 21-24 kDa (BERGH et al., 1985; MACFARLANE, 1999). Em

adição à CP, o RNA 2 pode codificar de uma a três proteínas não estruturais (9k, 2b e 2c) as quais, de acordo com estudos de vírus pseudo-recombinantes e mutantes, têm sido relacionadas com a transmissão dos vírus por nematóides (PLOEG et al., 1993; MACFARLANE et al., 1996; HERNÁNDEZ et al., 1997; VELLIOS et al., 2002). O sequenciamento de mais de doze isolados de tobravírus revelaram uma grande diversidade na sequência primária de nucleotídeos e no conteúdo gênico do RNA 2. Geralmente essas variações são resultantes da recombinação entre regiões não-homólogas dos RNAs 1 e 2 permitindo a troca de sequências da extremidade 3' do RNA 2 com aquelas do RNA 1 (MACFARLANE, 1999).

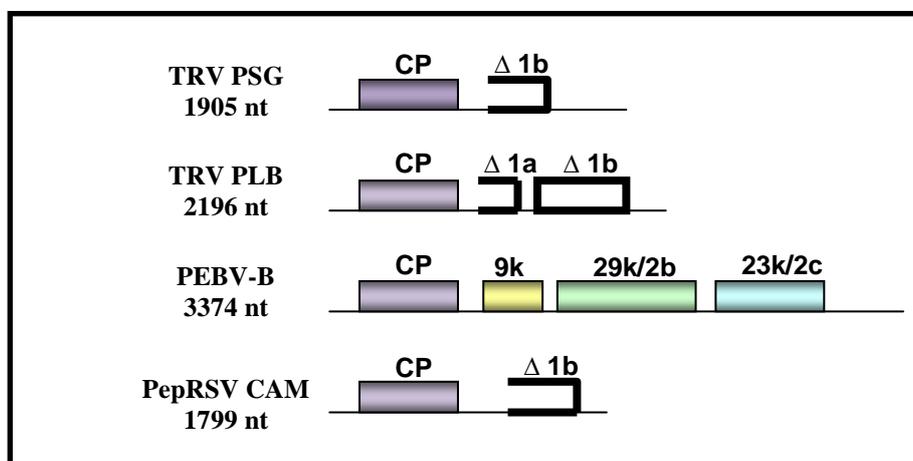
Os dois componentes genômicos (RNA 1 e RNA 2) são protegidos na região 5'-terminal por uma estrutura conhecida como “cap”, não são poliadenilados na região 3'-terminal, mas possuem uma estrutura semelhante ao RNA transportador nessa extremidade (VAN BELKUM et al., 1987).

Uma característica particular dos tobravírus é a habilidade de causar dois tipos de infecção. Os isolados que contêm os dois tipos de RNAs são denominados “tipo-M” (ou multiplicadores) e apresentam capsídeo, uma vez que o RNA 2 contém o gene que codifica a proteína capsidial. O isolado viral “tipo-NM” (ou não-multiplicador) contém apenas o RNA 1 e pode iniciar, por si só, o processo de infecção na hospedeira, indicando conter informação genética para sua replicação e transporte intercelular. Porém, neste caso, não há produção de proteína capsidial (MACFARLANE, 1999). Embora, a descrição inicial dos isolados NM como “não-multiplicadores” seja, portanto, errônea, as abreviações N ou NM continuam sendo utilizadas por estarem muito bem estabelecidas (SWANSON et al., 2002).

Inicialmente acreditava-se que os isolados tipo-NM se espalhavam pela planta apenas por movimento célula a célula e não, via floema (HARRISON e ROBINSON, 1986). Experimentos recentes mostraram que plantas inoculadas com isolados tipo-NM de TRV e PEBV apresentaram sintomas sistêmicos após 4-6 dias. Esse rápido espalhamento do RNA 1 pela planta sugere a utilização dos vasos do floema para movimentação, uma vez que a passagem apenas via célula a célula levaria semanas para ocasionar os sintomas observados (SWANTSON et al., 2002).



**Figura 3.** Organização genômica do RNA 1 dos tobravírus. Cada ORF está representada por um retângulo. O tamanho das proteínas em kDa está indicado por K e os nomes aparecem sobre cada gene. H, domínio helicase; Rep, Helicase/RdRp; 29-30K, provável MP; 12-16K, proteína, provavelmente, relacionada com a transmissão do vírus por sementes. As setas representam o códon de terminação tipo “opal”. (Figura Adaptada de MACFARLANE, 1999).



**Figura 4.** Organização genômica do RNA 2 dos tobravírus. CP, gene que codifica para a proteína capsial. Os retângulos destacados com negrito representam regiões derivadas do RNA 1 por recombinação, sendo que, retângulos abertos denotam cópias incompletas de genes do RNA 1. (Figura adaptada de MACFARLANE, 1999).

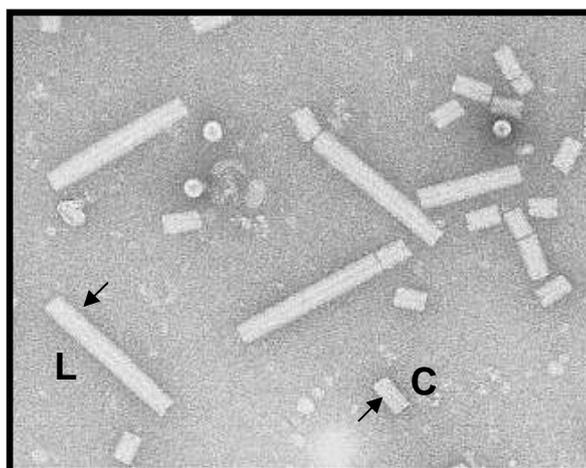
Os tobravírus são transmitidos naturalmente por nematóides dos gêneros *Trichodorus* e *Paratrichodorus* de maneira não-circulativa, ou seja, o vírus fica restrito ao aparelho bucal do agente transmissor (HARRISON e ROBINSON, 1986; VISSER e BOL, 1999; MACFARLANE e BROWN, 1995).

Tanto a proteína capsidial quanto as outras três proteínas (9k, 2b e 2c) codificadas pelo RNA 2 do PEBV têm sido associadas com a transmissão viral por nematóides (MACFARLANE et al., 1996).

### 1.2.1 *Pepper ringspot virus (PepRSV)*

O *Pepper ringspot virus* foi encontrado no Brasil pela primeira vez infectando pimentão (*Capsicum annuum* L.) sendo caracterizado por COSTA et al. (1961) como integrante do gênero *Tobravirus*. Durante muitos anos o PepRSV foi conhecido como isolado CAM do TRV, mas hoje é considerado uma espécie à parte (HARRISON e ROBINSON, 1986). No Brasil a denominação “Vírus do Anel do Pimentão” (VAP) foi amplamente utilizada devido aos típicos sintomas em forma de anel causados pelos vírus, porém essa nomenclatura está em desuso.

As partículas longas do PepRSV possuem comprimento de 197-200 nm e as curtas 52-55 nm, ambas com diâmetro da ordem de 20 nm (Figura 5) (KITAJIMA, 1967; KITAJIMA et al., 1969). As partículas virais são formadas por 95% de proteína e 5% de RNA (SILVA, 1965).



**Figura 5.** Micrografia eletrônica de transmissão mostrando partículas longas (L) e curtas (C) do PepRSV. Aumento de 118.000 X (*Micrografia eletrônica obtida de GASPAR et al., 1984*)

SILVA (1965) e OLIVEIRA (1967) verificaram que, assim como nas outras espécies do gênero, as partículas virais longas do PepRSV são infectivas por si só, ou seja, carregam a informação genética para um sistema RNA-replicase e partículas curtas não são infectivas isoladamente, porém, quando uma mistura dos dois tipos de partículas é usada como inóculo, há infecção com a produção de partículas virais completas (vírions).

O RNA 1 do PepRSV possui 6,8 Kb e codifica quatro proteínas distintas (Figura 3 c) (BERGH e SIEGEL, 1989; MACFARLANE, 1999). Duas delas (massas moleculares 136 e 195 kDa) são traduzidas a partir da extremidade 5' sendo que a proteína de 195 kDa (Replicase) é produzida por tradução através do códon de terminação (UGA) da proteína de 136 kDa (Helicase) (FRITSCH et al., 1977; PELHAM, 1979). Dois outros polipeptídeos de peso molecular 29 kDa e 12 kDa são também codificados pelo RNA 1 mas são traduzidos a partir de RNAs sub-genômicos de 1,5 Kb e 0,8 Kb respectivamente. Esses dois RNAs sub-genômicos não são protegidos por proteína capsidial diferentemente do que ocorre com vários isolados do gênero onde o RNA sub-genômico de 1,5 Kb é encapsidado (BERGH e SIEGEL, 1989).

O RNA 2 possui 1,8 Kb e codifica apenas um polipeptídeo de massa molecular 24 kDa identificado como a proteína capsidial (MAYO et al., 1976). Esta única ORF é flanqueada, nas extremidades 5' e 3', por regiões não codificadoras e de função desconhecida. A região terminal 3' do RNA 2 tem sequência de 459 nt idêntica à parte da região terminal 3' do RNA 1 (BERGH et al., 1985; MACFARLANE, 1999).

Tecidos diversos de diferentes plantas hospedeiras infectadas pelo PepRSV mostram agregados de partículas virais de 200 nm de comprimento dispostas paralelamente no citoplasma, em estreita associação com mitocôndrias (Figura 6) (KITAJIMA, 1967; HARRISON e ROBERTS, 1968; KITAJIMA e COSTA, 1969). Esta é uma característica específica do PepRSV, uma vez que nenhuma das outras duas espécies do gênero apresenta associação de partículas com mitocôndrias. As partículas curtas usualmente não estão associadas às mitocôndrias, sendo encontradas livres no citoplasma. Entretanto, agregados de partículas curtas, também associados a mitocôndrias, foram encontrados em grãos de pólen de tomateiro (CAMARGO et al, 1969),

em tecido foliar de fumo (NOGUEIRA, 1978) e em tecido foliar e de antera de tomateiro (GASPAR et al., 1984). Ainda não é conhecido se essa associação entre o PepRSV e mitocôndrias desempenha algum papel no ciclo de multiplicação do vírus.

NOGUEIRA (1978), através de estudos autorradiográficos ao microscópio eletrônico, identificou que o principal local de biossíntese do RNA viral é o nucléolo da célula infectada, embora alguma síntese possa ocorrer no nucleoplasma, entretanto esses dados carecem de confirmação.



**Figura 6.** Micrografia eletrônica de transmissão mostrando corte de célula esporogênica de antera de tomateiro apresentando partículas longas (L) e curtas (C) do PepRSV associadas com mitocôndrias (M). N, núcleo; TAP, tapete. (Figura obtida de GASPAR et al., 1984).

### 1.3 Síntese das proteínas virais durante a infecção

A replicação dos fitovírus portadores de RNAs mensageiros de fita simples (aproximadamente 90% dos vírus de plantas) pode ser resumida como segue. Após a entrada do vírus na célula (através de "ferimentos" ou por ação de vetores) a partícula viral deve ser desmontada, isto é, deve haver imediata separação do RNA das subunidades proteicas da capa.

Os ribossomos movem-se então ao longo do RNA viral removendo a capa proteica e, ao mesmo tempo, traduzindo o primeiro códon próximo a região 5'. A tradução deste códon produz uma RNA polimerase viral que, provavelmente, torna-se associada à proteína(s) da hospedeira, formando

um "complexo de replicação", experimentalmente evidenciado para vários vírus de plantas (HULL, 2004).

O complexo de replicação usa a fita de RNA com polaridade positiva como um molde para produzir, a partir da região 3', cópias complementares de RNA com polaridade negativa, via formação de estruturas conhecidas como "formas replicativas" (RFs) (HULL, 2002).

O dogma de tradução dos ribossomos eucarióticos determina que, em um RNA mensageiro policistrônico, somente o códon da região 5' é traduzido, sendo que qualquer códon que esteja localizado abaixo dessa região (portanto em direção a região 3') não é traduzido. Para contornar esse problema, os vírus de plantas empregam estratégias distintas para a expressão desses genes, incluindo-se a produção de poliproteínas, a segmentação do genoma, a síntese de RNAs subgenômicos, a mudança de fase ("frameshift") durante a tradução e outras (ZERBINI Jr. et al., 2002).

As subunidades proteicas da capa viral são produzidas em grande quantidade de modo que as fitas de RNA com polaridade positiva podem ser recobertas produzindo novas partículas virais, as quais acumulam, em geral, no citoplasma da célula (FIELDS, 1996).

Além da replicase e da proteína capsidial, os vírus codificam pelo menos uma proteína não-estrutural relacionada com movimento célula a célula das partículas. Em geral, a proteína do movimento é produzida em menores quantidades quando comparada à proteína capsidial. Alguns estudos mostram que essa proteína é rapidamente degradada após executar sua função (CARRINGTON et al., 1996).

Teoricamente, um vírus de planta necessita basicamente das três proteínas já citadas: uma replicase, uma proteína do movimento célula a célula e uma proteína capsidial. Muitos vírus, contudo, codificam outras proteínas com funções diversas como fatores de transmissão por afídeos, proteases, glicoproteínas, dentre outras. Em todas essas etapas, as proteínas codificadas pelo vírus devem interagir com proteínas do hospedeiro, a fim de executar suas funções e permitir o sucesso da infecção viral (HULL, 2002).

A regulação da expressão de proteínas determina a quantidade de proteína produzida ao longo do ciclo de vida do vírus ou permite a produção de proteínas apenas durante pontos específicos do ciclo. Muitas vezes está

intimamente relacionada com a estratégia de replicação do genoma e de expressão dos genes virais, incluindo segmentação do genoma, síntese de RNA subgenômicos, mudança de fase durante a tradução, entre outras (ZERBINI Jr. et al., 2002). A proteína capsidial também pode exercer importante papel no controle da expressão, como veremos a seguir.

Muitas proteínas codificadas pelos RNAs virais, sejam proteínas capsidiais ou proteínas não estruturais, acumulam-se na célula como "corpos de inclusão" ou "viroplasmas" tanto no citoplasma como no núcleo (AGRIOS, 1988).

Quando as proteínas virais acumulam nas células como corpos de inclusão há a possibilidade de as mesmas serem purificadas por métodos bioquímicos, a partir de extratos das folhas das plantas infectadas como ocorre, por exemplo, com os potyvírus (URCUQUI-INCHIMA et al., 2001). Por outro lado, quando as proteínas não estruturais (e mesmo as partículas virais) são sintetizadas em tecidos específicos da hospedeira (floema, por exemplo) e em baixas concentrações, tais métodos mostram-se ineficientes e alternativas devem ser buscadas. Uma estratégia alternativa que vem sendo utilizada ultimamente para estes casos é a expressão de tais proteínas, sejam estruturais ou não, em *E. coli*. Podendo, após purificação, serem utilizadas para a produção de antissoros específicos e para estudos de genômica estrutural (PACHECO et al., 2012).

### **1.3.1 Proteína capsidial – Uma proteína multifuncional**

Há algumas décadas acreditava-se que o papel da proteína capsidial era apenas o de envolver o genoma viral protegendo-o da degradação. Embora a montagem do vírion e sua dissociação continuem sendo a principal função dessas proteínas sabe-se hoje que as CPs exercem funções em, praticamente, todos os estágios do ciclo de infecção viral. As CPs caracterizam-se, portanto, como proteínas multifuncionais (IVANOV e MÄKINEN, 2012).

Assim, além da função de encapsidação da partícula viral, as proteínas capsidiais podem estar envolvidas com a transmissão por afídeos (ATREYA, 1995), movimento célula a célula (SCHOLTHOF, 2005), movimento

a longa distância (DOLJA et al., 1994), replicação viral (YI et al., 2009), determinação de sintomas, entre outros. Dessa forma, podemos dizer que, a proteína capsidial possui propriedades estruturais e não-estruturais (IVANOV e MÄKINEN, 2012).

Um dos papéis mais importantes da CP está relacionado ao controle da expressão de seu próprio gene. Como exemplo, podemos citar o que ocorre com o *Alfafa mosaic virus* (AMV), gênero *Alfamovirus*. Nestes vírus a CP é essencial para eficientes, replicação e tradução, do RNA viral em um fenômeno conhecido como 'ativação do genoma', onde a proteína capsidial se liga, simultaneamente, a extremidade 3' e ao cap 5', formando um 'looping' fechado. Considera-se, portanto, que a CP do AMV mimetiza a função da proteína de ligação de poli (A) (PABP) na tradução dos mRNAs (BOL, 2005).

Em alguns casos a CP não exerce papel direto na replicação do RNA, como por exemplo, no *Tobacco mosaic virus* (TMV), mas sua expressão permite que os complexos de replicação viral (CRVs) sejam formados e aumentem de tamanho mais rapidamente. Dessa forma, a CP do TMV exerce papel acessório na formação dos CRVs, o que possibilita, indiretamente, a regulação da produção dos RNAs subgenômicos (ASURMENDI et al., 2004).

As CPs também estão envolvidas no processo de espalhamento do vírus na planta hospedeira. A proteína capsidial exerce papel tanto no movimento célula a célula, que envolve uma variedade de mecanismos incluindo transporte pelos dos plasmodesmos através de túbulos e ligação à proteína do movimento (SCHOLTHOF, 2005).

Evidências genéticas mostraram que a função da CP do potyvírus *Tobacco etch virus* (TEV) vai além da formação de vírions. Mutantes na região central da CP do TEV não foram capazes de se movimentar célula a célula e, deleções na região amino-terminal da proteína também implicaram em limitação do movimento, o que mostra que essa proteína também pode atuar de forma acessória. Em outro experimento, verificou-se o envolvimento da CP na movimentação a longas distâncias. Mutantes de TEV apresentando deleções nas porções amino e carboxiterminais da CP, não puderam se movimentar sistemicamente na planta hospedeira (DOLJA et al., 1995).

A CP também pode ser utilizada para fins diagnósticos, na produção de antissoros. A produção de antissoros específicos para a detecção

de vírus geralmente utiliza como antígeno o vírus purificado. Porém, em alguns casos, fatores como baixo título viral na planta e instabilidade das partículas, dificultam sua purificação, caracterização e detecção. Quando não é possível obter as partículas virais purificadas, utilizam-se como imunógenos a CP e outras proteínas virais expressas *in vitro* (ZERBINI Jr. et al., 2002). A utilização da CP para a produção de antissoros é mais frequente que de outras proteínas virais não estruturais. Isso se dá, pois, a proteína capsidial é expressa em maior quantidade na planta infectada, facilitando sua detecção.

Com a utilização de técnicas de biologia molecular, tem se tornado comum a clonagem do gene da proteína capsidial de vírus em vetores que permitem sua expressão em *E. coli*, uma vez que para a grande maioria dos vírus de plantas, a proteína capsidial costuma ser a mais imunogênica. Desta forma, a produção de um antissoro produzido a partir da proteína expressa *in vitro* apresentará as mesmas propriedades daquele produzido a partir das partículas virais purificadas (BARBIERI et al., 2004).

Outro papel importante atribuído a CP é a iniciação do processo de auto-montagem (“self-assembly”) da partícula viral. Estudos realizados com a CP do TMV mostraram que ocorre uma reação de reconhecimento entre a CP e o RNA viral que dá início a montagem e alongação das partículas (STUBBS, 1997).

Esse fenômeno foi mostrado pela primeira vez em estudos com o TMV, onde os capsídeos se dissociavam com o aumento do pH da solução, mas eram reversivelmente rearranjados quando o pH era reduzido as condições iniciais (FRAENKEL-CONRAT e WILLIAMS, 1955).

Embora nas células hospedeiras a montagem das partículas virais seja mais complexa, envolvendo o RNA viral, proteínas não-estruturais e outros constituintes, há um bom número de vírus onde a montagem *in vitro* pode ocorrer na presença, apenas, de partículas purificadas. Exemplos bem estudados de vírus icosaédricos como o *Cowpea chlorotic mottle virus* (CCMV), o Vírus da hepatite B (HBV) e o Papiloma vírus humano (HPV) sugerem que todas as informações necessárias para a auto-montagem das partículas são codificadas pelos próprios capsômeros (JOHNSTON et al., 2010).

### 1.3.2 O movimento dos vírus nas plantas e a proteína do movimento

Os fitovírus compartilham muitas características com os vírus de animais, dentre elas: o modo de transmissão por vetores invertebrados; morfologia e composição das partículas e estratégias para expressão e replicação gênica. Entretanto, existe uma diferença fundamental, visto que para realizar o movimento de uma célula para outra adjacente, os vírus de planta precisam atravessar uma barreira substancial, a parede celular. A entrada inicial dos vírus nas células das plantas ocorre por penetração física da parede celular, por exemplo, por injúria mecânica de células epidérmicas, por insetos, ácaros, fungos ou nematoides que, ao se alimentarem podem transmitir o vírus de uma planta infectada para outra sadia. O transporte subsequente do material viral de uma célula inicialmente infectada para uma célula adjacente ocorre através dos plasmodesmos (KITAJIMA e LAURITIS, 1969).

Os plasmodesmos são canais estreitos revestidos pela membrana plasmática e atravessados por um túbulo de retículo endoplasmático modificado, conhecido como desmotúbulo. Esses canais permitem a continuidade do citoplasma e do sistema de endomembranas entre as células (RAVEN et al., 2001). Em células vegetais sadias, apenas moléculas com massa molecular inferior a 1 kDa são capazes de atravessar passivamente esses canais, que possuem limite de exclusão muito inferior à massa molecular dos menores fitovírus, ou mesmo de seus RNAs ou DNAs genômicos (WOLF et al., 1989). Então, como os fitovírus conseguem atravessar um canal tão estreito?

Para suprir essa restrição física, os fitovírus codificam proteínas especializadas, denominadas Proteínas do Movimento (MPs), capazes de dilatar os plasmodesmos permitindo a passagem de complexos formados por proteínas mais ácidos nucléicos virais ou mesmo de vírions inteiros (DEOM et al., 1992).

Os fitovírus são classificados em três grupos principais com base na necessidade ou não da proteína capsidial (CP) para atravessarem os plasmodesmos (Tabela 1). O grupo I compreende os vírus que não necessitam da CP para realizarem o movimento. Nos vírus do grupo II a CP age auxiliando a MP no movimento e protegendo o genoma viral. Os vírus do grupo III também

requerem a CP, porém, são transportados como partículas inteiras (SCHOLTHOF, 2005).

**Tabela 1.** Classificação dos fitovírus com base na necessidade da proteína capsidial para o movimento.

<b>Classificação</b>	<b>Proteína Capsidial (CP)</b>	<b>Número de MPs</b>	<b>Exemplo (gênero)</b>
Grupo I	Não é necessária	1	<i>Tobamovirus</i>
		2	<i>Carmovirus</i>
		3	<i>Hordeivirus</i>
Grupo II	Necessária (não atravessam como partículas)	1 (podem formar túbulos)	<i>Cucumovirus</i>
		2	<i>Potyvirus</i>
		3	<i>Potexvirus</i>
		1-2	<i>Geminivirus</i>
		1 (formam túbulos)	<i>Tospovirus</i>
Grupo III	Necessária (atravessam como partículas)	1 (formam túbulos)	<i>Comovirus</i>
		3 - 4 (e 2 CPs)	<i>Closterovirus</i>

(Tabela adaptada de SCHOLTHOF, 2005.)

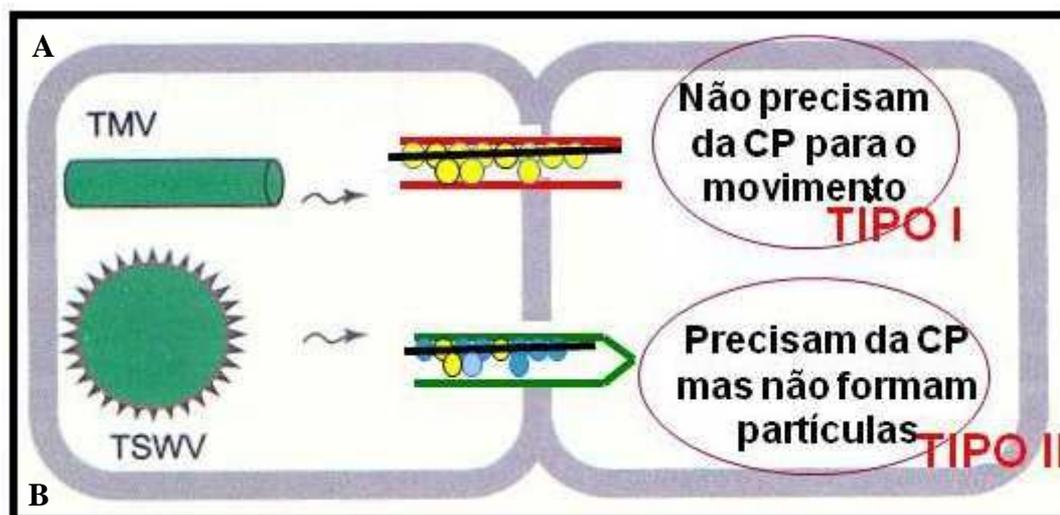
O TMV é membro representativo do grupo I. A atividade da MP do TMV em plasmodesmos, foi primeiramente demonstrada por WOLF et al., (1989), usando técnicas de microinjeção de dextranas fluorescentes de tamanhos variados em plantas transgênicas que expressavam a proteína de 30 kDa do TMV, a MP. Essas plantas apresentaram limite de exclusão dos plasmodesmos dez vezes maior do que o das plantas controle. Esse aumento permite o transporte dos complexos ribonucleoproteicos (VNPs) de uma célula adjacente à outra (Figura 7A) (CARRINGTON et al., 1996). A princípio, foi sugerido que além da interação com os plasmodesmos, a MP também apresentaria interações com os microtúbulos e filamentos de actina realizando um papel importante no direcionamento do complexo genoma-MP até o plasmodesmo (HEINLEIN et al., 1995; MCLEAN et al., 1995). Entretanto, alguns autores acreditam que esta associação está relacionada com a degradação da MP durante estágios avançados da infecção (PADGETT et al.,

1996; TZFIRA et al., 2000). Além disso, foi demonstrado que a ruptura de microtúbulos por agentes farmacológicos ou por silenciamento do gene da alfa-tubulina, não interfere significativamente no espalhamento do TMV pela planta (GILLESPIE et al., 2002). Outros exemplos de fitovírus que utilizam o movimento tipo I são os carmovírus e os hordeivírus que codificam duas e três MPs respectivamente (HULL, 2002; MOROZOV e SOLOVYEV, 2003).

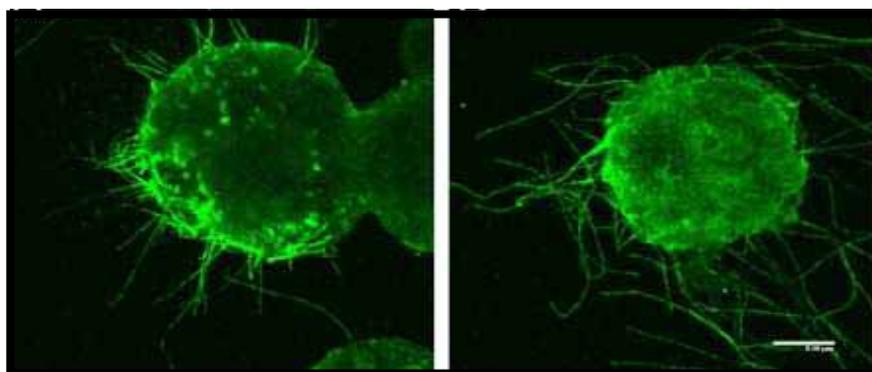
Os vírus do grupo II, assim como os do grupo I, codificam uma ou mais MPs. O *Cucumber mosaic virus* (CMV), por exemplo, codifica uma MP que compartilha muitas características bioquímicas com a MP do TMV, exceto que, a CP do CMV é necessária para o movimento célula a célula (SCHOLTHOF, 2005). NAGANO et al. (2001) mostraram que o requerimento da proteína capsidial para o movimento está relacionado com a porção carboxi-terminal da MP. Outro exemplo é o grupo dos geminivírus que, além das proteínas relacionadas com o movimento célula a célula, apresentam proteínas especializadas para a saída do DNA da região de replicação desses vírus (núcleo celular). Após deixar o núcleo, o DNA vai formar um complexo envolvendo duas MPs e a CP para o movimento através dos plasmodesmos (NOUEIRY et al., 1994). Os tospovírus, como o *Tomato spotted wilt virus* (TSWV) utilizam uma nucleoproteína N, análoga a CP, para o movimento e, por isso, pertencem ao grupo II (Figura 7B) (STORMS et al., 1995). Entretanto, o TSWV induz a formação de túbulos para o movimento intercelular. O CMV é, normalmente, transportado como complexo ribonucleoproteico, mas, é também, capaz de formar túbulos (SCHOLTHOF, 2005).

O movimento através de túbulos (movimento “túbulo-guiado”) é utilizado por uma grande variedade de vírus de planta, dentre eles: tospovírus, comovírus, caulimovírus, bromovírus, alfamovírus (LAZAROWITZ e BEACHY, 1999; CARRINGTON et al., 1996). Nesses casos, os plasmodesmos das células infectadas são drasticamente modificados, os desmotúbulos são removidos e um túbulo formado por MP é inserido no poro plasmodesmal. É por esse túbulo que as partículas virais são transportadas, sendo que neste tipo de transporte a proteína capsidial é também requerida (LAZAROWITZ e BEACHY, 1999; CARVALHO et al., 2004). Uma característica notável da MP dos caulimovírus, comovírus e tospovírus é sua habilidade de formar túbulos em protoplastos e em culturas de células de insetos (Figura 8). Para ambos os

sistemas a expressão da MP resulta em protusões de 50  $\mu\text{m}$  de comprimento extendendo-se pela superfície celular. A membrana plasmática é contínua com o interior dos túbulos (VAN LENT et al., 1991; CARRINGTON et al., 1996).



**Figura 7.** A) Movimento célula a célula do TMV e B) do TSWV. As esferas amarelas representam a proteína do movimento e as esferas azuis representam a proteína capsidial, ambas associadas ao RNA viral (linha preta). A linha vermelha representa o desmotúbulo e a linha verde o túbulo formado pelo vírus. (Figura adaptada de SCHOLTHOF, 2005).



**Figura 8.** Formação de túbulos em protoplastos infectados com TMV. As figuras mostram a expressão de proteínas relacionadas ao movimento dos vírus fusionadas a uma proteína imunofluorescente (GFP) (Figura obtida de LEWANDOWSKI e ADKINS, 2005).

Os vírus do grupo III também requerem a CP para o transporte, porém, neste caso as partículas virais atravessam os plasmodesmos. Os comovírus, por exemplo, induzem a formação de um túbulo por onde as partículas esféricas são transportadas para a célula adjacente (POUWELS et al., 2004). Já os membros da família *Closteroviridae* não induzem a formação

de túbulos, sendo transportados como partículas por um aumento da abertura dos plasmodesmos (ALZHANOVA et al., 2004).

As MPs virais podem ser agrupadas em quatro superfamílias: I) MPs codificadas pelo bloco triplo de genes dos potexvírus e vírus relacionados; II) MPs dos tymovírus; III) uma série de pequenos polipeptídeos menores que 10 kDa, codificados pelos carmovírus e alguns geminivírus; IV) a superfamília '30 K', relacionada com a MP de 30 kDa do TMV. Para a superfamília '30 K' foram propostos dois modelos de movimento, o primeiro é exemplificado pelo TMV e está relacionado com o aumento do limite de exclusão dos plasmodesmos e o segundo, típico dos comovírus, ocorre através de túbulos. MELCHER (2000) após alinhar a sequência de aminoácidos das MPs de 30 kDa de vários gêneros de fitovírus mostrou que os agrupamentos formados estão relacionados com o tipo de movimento realizado pelo vírus. Um grupo reúne os vírus formadores de túbulos e outro, aqueles vírus que não se movimentam por túbulos. O gênero *Tobravirus* encontra-se no segundo grupo.

O movimento a longa distância ou floema-dependente requer que o vírus atravesse as células do mesófilo, atingindo as células companheiras e, finalmente, entre nos elementos de tubos crivados, sempre via plasmodesmos (CARRINGTON et al., 1996). Mesmo aqueles vírus que se movem célula a célula independentemente da proteína capsidial, necessitam dessa para o movimento a longa distância. Mutantes de TMV apresentando modificações na proteína capsidial, por exemplo, movem-se célula a célula eficientemente, mas apresentam deficiência no movimento a longa distância (DAWSON et al., 1988). A forma pela qual a proteína capsidial auxilia no movimento sistêmico ainda não foi esclarecida, bem como muitos aspectos relacionados ao movimento dos vírus célula a célula permanecem mal compreendidos.

### 1.3.3 Helicases

Conforme descrevemos no item 1.3 a maioria dos vírus de plantas possui genoma composto por RNA de fita simples (ssRNA). Como as plantas não possuem enzimas capazes de replicar a molécula de RNA viral, os próprios vírus precisam codificar uma enzima para essa função, as polimerases de RNA dependente de RNA ou RdRps. As replicases já foram caracterizadas

para muitos vírus de plantas e, na maioria dos casos, existem duas funções principais: helicase e polimerase, que podem estar presentes em duas proteínas distintas. Cada uma dessas funções está associada a sequências específicas de aminoácidos. Atualmente uma proteína pode ser designada helicase ou polimerase, com base, simplesmente, na presença dessas regiões (HULL, 2002).

Helicases são proteínas motoras que utilizam a energia livre liberada da hidrólise dos nucleotídeos trifosfatos (NTPs) para separar as fitas duplas de DNA e RNA através da ruptura das pontes de hidrogênio. Elas são fundamentais para a replicação gênica, transcrição, tradução, recombinação e reparo (ABDELHALEEM, 2010).

De acordo com a presença de regiões com sequências conservadas de aminoácidos, as helicases podem ser classificadas em seis superfamílias (GORBALENYA e KOONIN, 1993; SINGLETON et al., 2007). As helicases de fitovírus estão agrupadas em apenas três das superfamílias: SFI, SFII e SFIII. As superfamílias I e II possuem sete regiões conservadas enquanto a superfamília III possui apenas três (CORDIN et al., 2006). O PepRSV, assim como a maioria dos vírus de planta com ssRNA sentido positivo, possui helicase pertencente a superfamília I, incluindo os furovírus, closterovírus, bromovírus, tobamovírus, dentre outros (KADARÉ e HAENNI, 1997).

Trabalhos com proteínas helicases geralmente estudam sua atividade enzimática como ATPase e RNA helicase. JIN e PETERSON (1995), investigaram a região carboxi-terminal da proteína não-estrutural 3 (NS3) do *Hepatitis C virus* (HCV), possivelmente envolvida com a função ATPase/Helicase. Para tal, promoveram a expressão *in vitro* e a purificação da proteína para medir sua atividade enzimática. Foi identificado que a região estudada possui ambas atividades NTPase e RNA helicase. Também verificou-se que ATP é o substrato preferido para a atividade ATPase e que a atividade RNA helicase depende de ATP e cátions divalentes, podendo ser utilizado o magnésio.

Estudos realizados por FERNÁNDEZ et al. (1995) para avaliar o envolvimento da Proteína CI (putativa helicase) do *Plum pox virus* (PPV) no processo de replicação viral e sua participação no movimento célula a célula durante a infecção viral, mostraram que a proteína expressa *in vitro* por meio de células bacterianas era extremamente insolúvel. Vários métodos de solubilização e renaturação testados não foram eficientes para obter quantidade razoáveis da proteína solúvel. Optaram, então, pela expressão da proteína fusionada a maltose e sua purificação em resina contendo amilose. Nesse caso as perdas da proteína CI nos processos de lavagens foi muito grande e a quantidade de proteína eluída não foi suficiente para a caracterização enzimática.

Dessa forma, a expressão *in vitro* do fragmento da helicase do PepRSV é possível de ser realizada, porém, não descartamos a existência de dificuldades na etapa de purificação e na obtenção da proteína solúvel. Dessa forma, a caracterização estrutural seria um terceiro desafio, devido a grande probabilidade de a proteína estar na fração insolúvel. Por outro lado, se quantidade suficientes de helicase em estado solúvel for obtida, será possível obter maiores informações sobre as helicases da superfamília I.

Estruturas tridimensionais de helicase já foram obtidas para o Hepatite C virus e para o Dengue vírus (DENG), ambos pertencentes a superfamília II das helicases (LOVE et al., 1996; KIM et al., 1998; MURTHY et al., 1999). Estruturas de cristais de helicases da superfamília III também já foram determinadas, e seus mecanismos estruturais descritos em termos estruturais (HICKMAN e DYDA, 2005). Entretanto, embora a helicase do grupo SFI de *Bacillus stearothermophilus* e *Escherichia coli* tenham sido estudada e, apresentado um sítio catalítico semelhante ao REC A do HCV que pertence ao grupo SFII (HESSON et al, 1997), estruturas tridimensionais para helicases virais da superfamília I, ainda não foram obtidas (XIANG et al., 2012).

## **1.4 Expressão de proteínas heterólogas em bactérias**

Nas últimas duas décadas sistemas de expressão de proteínas em procariotos, particularmente em *E. coli*, têm sido explorados para a produção, em escala industrial, de uma variedade de proteínas com

propriedades terapêuticas, bem como, em pesquisas acadêmicas que envolvem a produção de proteínas heterólogas (PACHECO et al., 2012). As células procarióticas (*E. coli*) são normalmente preferidas para a expressão de proteínas heterólogas pelo fato de oferecerem algumas vantagens como (1) fonte de carbono de baixo custo para o crescimento, (2) rápido acúmulo de biomassa, (3) possibilidade de uso em processo de fermentação com alta densidade de células, (4) possibilidade de produção de proteínas em larga escala.

A utilização de *E. coli* para produção de proteínas heterólogas humanas tem sido relatadas com sucesso para cerca de 75% dos casos (PACHECO et al., 2012). Entretanto, a falta de componentes para modificações pós-traducionais e a produção de proteínas inativas devido à formação de corpos de inclusão, oferecem um significativo desafio no uso destes sistemas de expressão (GRASLUND et al., 2008; VILLAVERDE e CARRIO, 2003). Apesar disso, a habilidade das células procariotas acumularem até 80% de sua massa seca em proteínas recombinantes, ainda faz com que esse seja o sistema de expressão escolhido para uma ampla gama de aplicações (DEMAIN e VAISHNAV, 2009).

Um variável número de protocolos está à disposição, os quais descrevem várias estratégias para a conversão de proteínas inativas, expressas como corpos de inclusão insolúveis, em frações solúveis e biologicamente ativas (SAHDEV et al., 2008). De modo geral, estas estratégias podem ser subdivididas em: (1) procedimentos onde a proteína é reenovelada a partir dos corpos de inclusão (2) procedimentos, nos quais, fatores relacionados à expressão são modificados para que haja produção de proteínas solúveis (3) procedimentos, onde a proteína de interesse é obtida na fração solúvel através da produção de uma proteína de fusão (SAHDEV et al., 2008; SORENSEN e MORTENSEN, 2005).

Sabe-se que a purificação de proteínas expressas de forma solúvel pode ser feita em menor tempo, com baixos custos e com menos prejuízos à integridade das proteínas, em comparação com a purificação e renaturação a partir dos corpos de inclusão. Assim, muito tem sido feito no sentido de obter metodologias que possibilitem e otimizem a expressão de proteínas solúveis. Quando nenhuma das estratégias é eficiente e, as proteínas

continuam presentes apenas nos corpos de inclusão, inevitavelmente, as metodologias envolvendo o uso de agentes desnaturantes passam a ser utilizadas.

Uma das estratégias mais conhecidas para produção de proteínas recombinantes solúveis é a expressão a baixas temperaturas (SCHEIN, 1989). Sabe-se que as interações hidrofóbicas que levam a formação dos agregados protéicos são, altamente, favorecidas pelas altas temperaturas (KIEFHABER et al., 1991). Além disso, baixas temperaturas inibem a produção de “heat shock” proteases e aumentam a produção de chaperonas (FERRER et al., 2004).

Altos níveis de expressão de proteínas recombinantes em *E. coli* em geral resultam no acúmulo dessas proteínas *in vivo* como agregados insolúveis (KANE e HARTLEY, 1988; FAHNERT et al., 2004). Em geral, os corpos de inclusão são solubilizados pelo uso de altas concentrações de agentes desnaturantes tais como uréia e cloridrato de guanidina, juntos com um agente redutor como  $\beta$ -mercaptoetanol (CLARCK, 1998; LILIE et al., 1998). As proteínas solubilizadas dessa forma são então renaturadas pela lenta remoção do agente desnaturante na presença de agentes oxidantes (FISCHER et al., 1993). A solubilização das proteínas dos corpos de inclusão por agentes desnaturantes resulta na perda da estrutura secundária levando à exposição da superfície hidrofóbica (DILL e SHORTLE, 1991). Tal perda da estrutura secundária durante a solubilização e a interação entre as moléculas proteicas desnaturadas durante a renaturação, resultam na agregação dessas moléculas e são consideradas como a principal razão para a pobre recuperação de proteínas biologicamente ativas a partir dos corpos de inclusão, muitas vezes atingindo apenas 15-25% da produção total de proteínas (DATAR et al., 1993). Assim sendo, o maior desafio nesta área tem sido converter eficientemente estas proteínas insolúveis e inativas em produtos solúveis e corretamente renaturados (CLARCK, 2003; PANDA, 2003).

Há dois importantes aspectos a serem considerados na recuperação de proteínas biologicamente ativas a partir dos corpos de inclusão: a solubilização dos agregados proteicos e a renaturação da proteína solubilizada para obter as proteínas em uma forma biologicamente ativa. A renaturação das proteínas obtidas em estado desnaturante, a partir de corpos

de inclusão, é um processo que requer muitas etapas operacionais e, na maioria das vezes, resulta em recuperação muito baixa de proteína renaturada. A recuperação de proteínas biologicamente ativas a partir de corpos de inclusão envolve quatro etapas: (1) isolamento dos corpos de inclusão a partir de células de *E. coli*, (2) solubilização dos agregados de proteínas, (3) renaturação e (4) purificação da proteína solubilizada (LILIE et al., 1998; CLARK, 2003; VALLEJO e RINAS, 2004). Dentre estas etapas, a solubilização e a renaturação são as cruciais para a alta recuperação de proteínas biologicamente ativas. Os corpos de inclusão são geralmente separados dos demais componentes celulares usando-se centrifugação de baixa rotação após a lise das células bacterianas. Agregados semi-puros juntos com contaminantes são então solubilizados usando-se agentes desnaturantes em altas concentrações (6 - 8 M), tais como uréia ou cloridrato de guanidina (FISCHER et al., 1993).

O uso de uma matriz adequada, Sephacryl S-100, por exemplo, na cromatografia por exclusão de tamanho facilita a remoção simultânea do agente desnaturante, a renaturação da proteína recombinante e a separação das moléculas por tamanho (BATAS e CHAUDHURI, 1996). Como a cromatografia por exclusão de tamanho oferece múltiplas vantagens na troca de tampões, na renaturação de proteínas e na separação de monômeros dos agregados, ela se constitui em método ideal para o reenovelamento de proteínas, em altas concentrações, obtidas a partir de corpos de inclusão (FAHEY et al., 2000; GU et al., 2001).

As simultâneas, troca de tampão, renaturação e purificação de proteínas solubilizadas, podem ser feitas utilizando-se também a cromatografia de troca iônica, onde a proteína desnaturada se liga a uma matriz (SCHLEGL et al., 2003; LI e JANSON, 2004). Interações intermoleculares que levam à agregação são minimizadas, uma vez que as moléculas protéicas ficam isoladas através da ligação a uma matriz-suporte. O uso simultâneo de tampão livre de agente desnaturante e otimização das condições de eluição levam à purificação de proteínas em forma biologicamente ativas (YASHIMOTO e KUBIO, 1999; LI e SU, 2002).

## 1.5 Ferramentas de bioinformática aplicadas ao estudo de estrutura proteica

A bioinformática é uma ciência recente, que surgiu pela necessidade de analisar os dados biológicos contidos em sequências de nucleotídeos. Esta ciência une fundamentos multidisciplinares (biologia molecular, bioquímica, matemática, física) ao poder computacional para organizar, catalogar e calcular dados fragmentados obtidos experimentalmente. Atualmente, a bioinformática pode ser dividida em cinco grandes tópicos: genômica, transcrição e regulação, evolução e filogenia, predição de estrutura/função proteica e redes regulatórias (PEVZNER e SHAMIR, 2011).

Basicamente, a predição de estrutura e função proteica envolve as seguintes etapas: análise de similaridade entre sequências, alinhamento múltiplo, identificação e caracterização de domínios, predição de estrutura secundária, predição de acessibilidade por solvente, reconhecimento de dobramento, construção de modelos tridimensionais em detalhe atômico e por final, validação do modelo (EDWARDS e COTTAGE, 2003).

A confiabilidade da estrutura tridimensional construída depende da porcentagem de identidade da sequência da proteína-alvo com o modelo. Modelos de alta precisão são baseados em estrutura de proteínas com identidade de pelo menos 50% com a proteína-alvo, modelos de precisão média com 30-50% de identidade e de baixa precisão com identidade abaixo de 30% (LEVITT, 1999).

A aplicação de modelos estruturais de proteínas é extremamente ampla. A comparação de modelos de alta e média precisão pode ser utilizada para análise preditiva de função e interação com ligante (SANCHEZ et al., 1996; MATSUMOTO et al., 1995). Modelos de baixa precisão também podem ser úteis em comparações de estruturas. Embora a precisão e a confiabilidade do modelo com baixa identidade (<30%) seja significante menor do que modelos gerados por sequências com identidade alta (>30%), a topologia básica da proteína ou de um domínio específico ainda pode ser modelada com boa qualidade. Desta forma, estes modelos podem revelar relações funcionais

e estruturais entre proteínas que não são identificáveis pela sequência de aminoácidos (BAKER e SALI, 2001).

A análise estrutural e funcional de proteínas por automação apresenta ainda a vantagem de demandar menor recurso financeiro e de tempo (BAKER e SALI, 2001). As técnicas experimentais necessitam de reagentes e validação com repetição, o que requer aumento no investimento destes. Os dados obtidos experimentalmente podem ser os mesmos obtidos *in silico*, se bem desenvolvidos. Isto mostra que a predição *in silico* é uma importante ferramenta para uma análise rápida e de baixo custo, e por isso observamos a expansão rápida desta ciência nos últimos anos (PEVZNER e SHAMIR, 2011).

## 2. Justificativa

Os Carlavírus são um amplo grupo de abrangência global e afetam cultivo de plantas de grande importância econômica mundialmente, como por exemplo, a batata e o milho. Embora a descoberta dos vírus deste grupo não seja recente, dados sobre o ciclo da infecção ainda não são totalmente elucidados. Consequentemente, não há controle totalmente eficiente contra os vírus do grupo.

O entendimento das relações do complexo vetor-vírus-hospedeira é uma premissa básica para traçar estratégias de controle integrado das viroses vegetais. Como parte desses esforços, importantes contribuições quanto à caracterização molecular dos vírus de plantas têm sido feitas citando-se, como exemplos, o sequenciamento dos RNAs, a detecção dos RNAs subgenômicos e a determinação da estratégia para expressão do genoma do vírus. Entretanto, muitos aspectos a respeito do ciclo de multiplicação dos vírus e da função de algumas proteínas não são conhecidos. Há carência de informações essenciais sobre o processo de infecção e replicação nos tecidos das plantas, o espalhamento célula a célula e sistêmico do vírus, bem como, a relação do vírus com o vetor. Estas informações são fundamentais para entender melhor a epidemiologia das doenças e estabelecer estratégias mais eficientes de controle.

Para os gêneros *Carlavirus* e *Tobravirus*, sabe-se que as proteínas do movimento, a helicase e a proteína capsidial são essenciais para o sucesso da infecção e replicação na planta. Entretanto, a estrutura destas proteínas não é conhecida em todos os indivíduos do grupo, o que dificulta o entendimento das diferenças do ciclo viral que este grupo apresenta.

Estudos de cristalização e análise por difração de raios-X, que pudessem fornecer dados sobre a estrutura das proteínas virais, são escassos, inclusive para a proteína capsidial. Isso provavelmente ocorre devido ao fato de que os vírus isométricos permitem a formação de cristais verdadeiros enquanto que partículas virais alongadas e flexuosas em geral formam cristais líquidos, nos quais as partículas são arranjadas regularmente em duas dimensões. Assim, análises por difração de raios-X não podem ser aplicadas no estudo de partículas com morfologia alongada. Nesse sentido, a cristalização e análise

por raios-X de proteínas capsidiais purificadas (não como partículas) e de proteínas não estruturais (helicases, polimerases, movimento, transmissão por vetores, etc) poderiam fornecer dados importantes sobre a estrutura dessas proteínas e de suas possíveis funções no ciclo de multiplicação viral.

O atraso na busca dessas informações deve-se fundamentalmente as dificuldades para obtenção de quantidades razoáveis de proteínas estruturais e, principalmente, as não-estruturais a partir de plantas infectadas. Como parte das estratégias para avançar nessa linha de pesquisa procurou-se expressar *in vitro* algumas proteínas essenciais ao ciclo de vida do CoLV e PepRSV e utilizá-las para estudos de caracterização molecular e estrutural.

Aliadas ao estudo estrutural experimental, análises de bioinformática auxiliam na organização de dados e em *insights* sobre a estrutura e a função. Estas ferramentas não são financeiramente custosas e apresentam boa confiabilidade.

Desta forma, os resultados estruturais das proteínas do capsídeo, do movimento e da helicase obtidas neste estudo poderão contribuir para a montagem do ciclo viral deste grupo e de estratégias para o controle de plantas infectadas por este e para utilização em biologia molecular na manipulação do DNA vegetal.

### 3. Objetivos

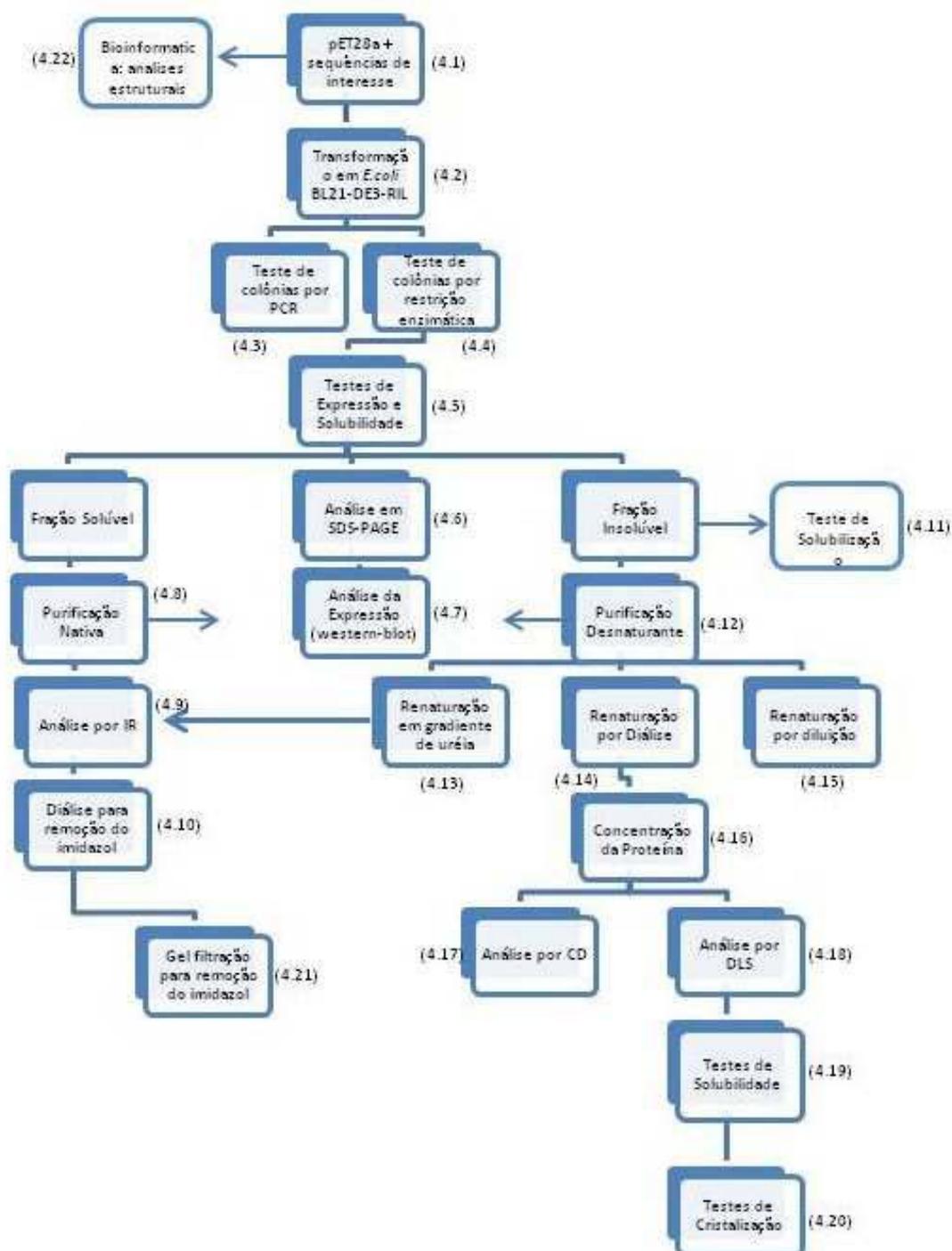
Este trabalho teve como objetivo geral a produção de proteínas em estado nativo, a partir de expressão em *Escherichia coli*, em quantidades e condições adequadas para a realização de estudos moleculares e estruturais dessas moléculas. Os seguintes vírus e proteínas foram estudados:

- Proteína Capsidial (CP) do *Cole latent virus*;
- Proteína do Movimento (MP) do *Pepper ringspot virus*;
- Um fragmento N-terminal da Proteína Helicase (HEL) do *Pepper ringspot virus*.

São objetivos específicos desse trabalho:

- Testes de expressão para verificar as melhores condições de produção de cada uma das proteínas;
- Obtenção das proteínas de interesse em estado nativo e excelentes condições de pureza;
- Análise do estado de agregação das proteínas obtidas;
- Estudos moleculares e estruturais (análises por dicroísmo circular, espectroscopia de infravermelho e espalhamento dinâmico de luz);
- Elaboração de modelos teóricos das proteínas de interesse;
- Comparação dos resultados experimentais com os modelos teóricos obtidos.

## 4. Material e métodos



**Quadro 1.** Fluxograma ilustrando a sequência de experimentos realizados no decorrer do trabalho. A numeração entre parênteses corresponde ao item descrito em material e métodos.

#### 4.1 Obtenção dos plasmídeos recombinantes

Os plasmídeos (pET28a/Novagen) Anexo 1 recombinantes (cedidos pelo Prof. José Osmar Gaspar – IBILCE/UNESP - São José do Rio Preto - SP) foram ponto de partida para os experimentos. Os oligonucleotídeos mostrados na Tabela 2 foram usados nas clonagens das sequências de interesse.

**Tabela 2.** Sequências dos oligonucleotídeos utilizados para amplificar os genes em estudo. Os sítios de restrição incorporados encontram-se discriminados em vermelho (**BamH I**) e azul (**Hind III**).

Gene	Oligo	pb	Sequências dos oligonucleotídeos
Capa proteica do <i>Cole latent virus</i> (CP-CoLV)	CP-CoLVexp (+)	921	5' <b>GGATCC</b> ATGGACCAAAGTAGAG 3'
	CP-CoLV exp (-)		5' <b>AAGCTT</b> TTTGCTACTTCACACACA 3'
Proteína do Movimento do <i>Pepper ringspot virus</i> (MP-PepRSV)	MP-PepRSVexp (+)	800	5' CTT <b>GGATCC</b> ATGGAGAACGA 3'
	MP-PepRSVexp (-)		5' AT <b>AAGCTT</b> AGAGCACACACTTCG 3'
Sequência parcial da Helicase do <i>Pepper ringspot virus</i> (Hel-PepRSV)	Hel-PepRSV exp (+)	845	5' <b>GGATCC</b> GAATCAGCAAACC 3'
	Hel-PepRSV exp (-)		5' <b>AAGCTT</b> CTCGACCGACACC 3'

#### 4.2. Transformação em células competentes

Os plasmídeos recombinantes foram transformados em linhagem BL21-DE3-RIL de *E. coli* competente utilizando-se 50 µL de células competentes e 1 µL de amostra de plasmídeo. Essa mistura permaneceu em gelo durante 30 min e, então, foi submetida ao choque de temperatura a 42 °C por 30 s e 2 min no gelo. Em seguida foram adicionados 450 µL de meio líquido SOC (meio SOB + 200 mM de glicose) e as bactérias foram submetidas à regeneração (1 h a 37 °C com agitação a 250 rpm). Posteriormente, as bactérias foram plaqueadas em meio sólido seletivo (LB + canamicina e cloranfenicol 100 µg/mL).

### 4.3 Confirmação da transformação por PCR

As colônias foram testadas por PCR, utilizando oligonucleotídeos específicos para cada uma das sequências das proteínas em questão (Tabela 2). A mistura da reação (25 µl) consistiu de: 20 mM Tris.HCl pH 9,5; 50 mM KCl; 5,0 mM MgCl<sub>2</sub>; 0,2 mM cada dCTP, dGTP, dATP, dTTP; uma pequena fração da colônia; 150 ng de cada oligonucleotídeo e 3 unidades de *Taq* polimerase (Fermentas). As reações ocorreram em Termociclador Mastercycler personal (Eppendorf) programado para desnaturação inicial a 94 °C por 4 min, seguida de 30 ciclos a 94 °C por 1 min, 45 °C por 1 min e 72 °C por 2 min, com uma extensão final a 72 °C por 10 min.

### 4.4 Confirmação da transformação por restrição enzimática

Uma das colônias positivas para a reação de PCR de cada inserto de interesse, foi repicada em 5 mL de meio LB contendo canamicina e cloranfenicol (100 µg/mL) e incubada *overnight* sob agitação de 250 rpm. Os plasmídeos foram extraídos utilizando o “GeneJet Plasmid Miniprep Kit”, seguindo-se as orientações do fabricante (Fermentas) e digeridos utilizando enzimas de restrição específicas para confirmar a presença do inserto.

Para cada 1 µg de DNA foram utilizadas 10 unidades (1µl) de cada enzima (BamHI e Hind III). A reação de corte de 20 µl consistiu de: 12 µl de vetor (125 ng/µl), 1 µl de cada enzima, 2 µl de tampão E 10X, 2,0 µl de BSA 10X e 2 µl de água. A mistura foi submetida à temperatura de 37 °C por 2h e a 65 °C por 15 min. Em seguida, a mistura foi submetida à eletroforese em gel de agarose 1%.

As colônias positivas foram utilizadas para os pré-inóculos nos testes de expressão.

### 4.5 Testes de expressão e solubilidade em diferentes temperaturas

Foram realizados testes de expressão em três temperaturas diferentes: 18°C, 25°C e 37°C. Para todas as temperaturas, foram testadas

duas concentrações de IPTG (0,5 mM e 1,0 mM) para indução. Os experimentos foram feitos como descrito a seguir.

Inicialmente 5 mL do pré-inóculo foram transferidos para um novo erlenmeyer contendo 50 mL de meio líquido seletivo (LB + canamicina e cloranfenicol 100 µg/mL). O tubo foi mantido sob agitação e aeração até que a  $DO_{550}$  atingisse entre 0,5 e 0,65 e, a seguir, a indução de expressão foi feita com IPTG. Foram coletadas amostras antes da indução, após 4 h e 18 h de indução.

As células foram recuperadas por centrifugação a 5000 g por 30 min, ressuspensas em 10 ml de tampão de lise não-desnaturante ( $NaH_2PO_4$  50 mM; NaCl 300 mM; Imidazol 20 mM; pH 7,4) e sonicadas. A amostra lisada foi centrifugada a 6.000 g por 30 min. O sobrenadante foi armazenado e o precipitado ressuspensado em 10 ml de tampão de lise. Essas amostras foram analisadas em SDS-PAGE para verificar a presença de proteínas de interesse nas frações solúvel e insolúvel.

#### **4.6 Análise em gel de poliacrilamida (SDS-PAGE)**

As amostras coletadas durante o teste de expressão foram analisadas através de eletroforese feita em gel desnaturante de poliacrilamida 12% (SDS-PAGE) segundo Laemmli (1970). As proteínas foram desnaturadas por fervura (3 min a 100 °C) em tampão de amostra (Tris.HCl 0,125 M, pH 6,8, contendo SDS 4%, glicerol 20%, 2-ME 10% e azul de bromofenol 1,0 mM). As amostras foram centrifugadas por 5 min a 16.000 x g e 10 µL do sobrenadante de cada amostra aplicados no gel. A eletroforese foi feita em aparelho Hoefer mini VE Amersham Biosciences GE. Os géis foram fixados e corados em mistura contendo metanol, água, ácido acético (50:50:10) e Comassie Blue 0,25% por 3 h e descorados com metanol e água (1:1) por 18h.

#### **4.7 Análise da expressão por western-blot com anti-poli-histidina**

Após separação em gel SDS-PAGE, as frações solúveis e insolúveis foram transferidas para membrana de nitrocelulose (Hybond-C Extra-Amersham Biosciences) em um eletrotransferidor (Mini Trans-blot

Electrophoretic Transfer Cell Bio-Rad) operando sob voltagem constante (100 V) por 60 min. O tampão de transferência utilizado foi o Tris.Glicina (Tris 48 mM, Glicina 39 mM, SDS 0.037%, metanol 20%). Em seguida, a membrana foi incubada por 1 h com o antissoro monoclonal anti-poli-histidina (produzido em camundongo; Sigma) na diluição de 1:3.000, lavada com tampão PBS-T e incubada por 1 h com antissoro anti-IgG de camundongo conjugado com fosfatase alcalina (Sigma) diluído 1:30.000 em tampão PBS-T contendo 0,3 % de BSA. Após lavagens com tampão PBS, a membrana foi incubada com o substrato 5-bromo-4-cloro-3-indolil-fosfato/nitro-blue tetrazolio (BCIP/NBT) até o aparecimento do resultado da reação (cor azul púrpura).

#### **4.8 Purificação das proteínas em condições nativas**

As proteínas solúveis presentes no sobrenadante foram purificadas em coluna de afinidade utilizando resina de matriz de níquel (Ni Sepharose High Performance) conforme instruções do fabricante (GE Healthcare). Após a ligação das proteínas, a lavagem foi feita com tampão ( $\text{NaH}_2\text{PO}_4$  50 mM; NaCl 500 mM e imidazol 40 mM). A eluição das proteínas foi feita em tampão ( $\text{NaH}_2\text{PO}_4$  50 mM; NaCl 300 mM e imidazol 250 mM, pH 7,4).

#### **4.9 Análise da CP do CoLV por espectroscopia de infravermelho (IR)**

As análises por infravermelho das amostras das proteínas (CP do CoLV purificada em condições nativas e CP do CoLV purificada em condições desnaturantes com posterior reenovelamento por gradiente de uréia) foram realizadas pelo Prof. Dr. Marinônio Lopes Cornélio, do Laboratório de Espectroscopia Fototérmica, Departamento de Física deste Instituto. Para cada amostra foram realizadas duas leituras do espectro, sendo a primeira com o tampão sem a proteína de interesse (branco) e a segunda o mesmo tampão contendo as proteínas de interesse. Essa subtração foi feita pelo software GRAMS/32<sup>®</sup>. O aparelho utilizado para as leituras foi o Espectrofotômetro de Nicolet Nexus 670 FT-IR com resolução de  $4 \text{ cm}^{-1}$  e um total de 256 interferogramas em cada coleta com a temperatura controlada a 25 °C com banho térmico (Thermo, NESLAB RTE 7).

#### **4.10 Diálise das frações eluídas para remoção do imidazol**

As frações eluídas foram transferidas para membrana de diálise (Fisherbrand- 6000-8000 MW) e mergulhadas, inicialmente, em tampão ( $\text{NaH}_2\text{PO}_4$ , NaCl 200 mM, pH 7,4) por 2 h, sob agitação e 20 °C. A seguir, o tampão foi substituído por outro semelhante contendo NaCl 100 mM e, após 2 h, por outro tampão contendo NaCl 50 mM.

#### **4.11 Testes de solubilização com agentes não-desnaturantes**

A fim de obter a proteína expressa na forma solúvel, foi feito um teste de expressão, de acordo com o item 4.2.5. Após a expressão, o meio foi separado em tubos diferentes e foi centrifugado por 30 min a 6.000 g e 4 °C. O “pellet” foi ressuspenso em diferentes tampões com agentes não desnaturantes. Os tampões utilizados, todos em pH 8,0 foram: 300 mM NaCl, 50 mM  $\text{NaH}_2\text{PO}_4$  e L-arginina 1 M; 300 mM NaCl, 50 mM  $\text{NaH}_2\text{PO}_4$  e sulfobetaina (0,5 M); 300 mM NaCl, 50 mM  $\text{NaH}_2\text{PO}_4$  e PEG 3300 (0,5 g/L); 300 mM NaCl, 50 mM  $\text{NaH}_2\text{PO}_4$  e ciclodextrina 100 mM. As soluções foram submetidas a pulsos de 15 s a 300 W em sonicador para lise celular. As amostras foram centrifugadas por 30 min a 6.000 g a 4 °C e alíquotas do sobrenadante de cada tubo foram retiradas para análise em gel de poliacrilamida (SDS – PAGE).

#### **4.12 Purificação das proteínas em condições desnaturantes**

As células recuperadas da etapa de expressão a 37°C foram ressuspenso em tampão de lavagem I (Tris.HCl 50mM pH 8,0 e 2% deoxicolato). A suspensão foi sonicada e centrifugada a 10.000 x g por 10 minutos. O sobrenadante foi descartado e o corpo de inclusão ressuspenso em tampão de lavagem II (Tris.HCl 50mM pH8,0 1% Triton e 10mM Mercaptoetanol). A suspensão foi sonicada e centrifugada a 10.000 x g por 10 minutos. Este processo foi repetido quatro vezes, alternando os tampões de lavagem.

A purificação do corpo de inclusão foi feita com tampões para purificação sob condições desnaturantes, em coluna de afinidade utilizando resina de matriz de níquel conforme instruções do fabricante. Foram feitas lavagens com tampões 100 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris-HCl e 8 M uréia e pH, respectivamente, 6,3 e 5,9. A eluição foi feita em tampão 100 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris.HCl e 8 M uréia, pH 4,5.

#### **4.13 Purificação das proteínas em condições desnaturantes e reenovelamento por gradiente de uréia**

As células recuperadas da etapa de expressão a 37°C foram submetidas ao procedimento descrito no item anterior.

A purificação do corpo de inclusão foi feita com tampões para purificação sob condições desnaturantes, em coluna de afinidade utilizando resina de matriz de níquel conforme instruções do fabricante (GE Healthcare). Após a ligação das proteínas, foi realizado o processo de reenovelamento em tampão 100 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris- HCl e gradiente 8 – 0 M uréia. A lavagem foi feita com tampão (NaH<sub>2</sub>PO<sub>4</sub> 50 mM, NaCl 500 mM e imidazol 40 mM). A eluição das proteínas foi feita em tampão (NaH<sub>2</sub>PO<sub>4</sub> 50 mM, NaCl 300 mM e Imidazol 250 mM, pH 7,4).

#### **4.14 Reenovelamento das proteínas por diluição**

Foram utilizados protocolos de diluição simples em tampão Tris.HCl pH 7,5 contendo 0,2 M NaCl, 1 mM DTT e 1 M NDSB201 [3(1-pyridinio)-1-propane sulfonate], a 4 °C, no qual solução com guanidina 6M contendo proteína de interesse (1 mg/mL) foi diluída até à concentração 10 – 100µg/ml. A solução contendo proteína purificada foi utilizada no experimento de purificação conforme o item 4.2.8.

#### **4.15 Reenovelamento das proteínas por diálise**

Foram utilizados protocolos de diluição simples em tampão Tris.HCl pH 7,5 contendo 0,2 M NaCl, 1 mM DTT e 1 M NDSB201 [3(1-

pyridinio)-1-propane sulfonate], a 4°C, no qual solução com guanidina 6 M contendo proteína de interesse (1 mg/mL) foi diluída até à concentração 10 – 100 µg/ml.

Também foram feitos experimentos de diálise em membrana de celulose com exclusão de 3.000 Da (Fisherbrand). As amostras de proteínas eluídas em tampão contendo 8 M de uréia, foram dialisadas contra 1 L de tampão fosfato/acetato pH 5,0 contendo 1% glicerol até que a concentração de equilíbrio chegasse a 0,01 M uréia. A diálise teve duração de 24 h com duas trocas de tampão.

#### **4.16 Concentração e quantificação das amostras purificadas**

Após a diálise, as amostras contendo a proteína purificada foram concentradas em dispositivos de ultrafiltração AMICON (3.000 Da) conforme instruções do fabricante (Millipore) a uma rotação de 4.000 g a 10 °C.

A quantificação das proteínas foi feita através da medida de absorbância das mesmas a 280 nm. Os espectros de absorbância foram obtidos em espectrofotômetro – Espectronic Genesys 2, com varredura de comprimento de onda de 340-200 nm. A concentração foi determinada a partir dos valores dos coeficientes de extinção molar ( $\epsilon$ ) obtidos através do programa ProtParam ([HTTP://ca.expasy.org](http://ca.expasy.org)) e dos valores de absorbância a 280 nm. A quantificação também foi feita através de leitura em equipamento Quant-iT™ Assays, conforme instruções do fabricante (Invitrogen).

#### **4.17 Análise das proteínas por Dicroísmo Circular (CD)**

Os espectros de Dicroísmo Circular (CD) foram obtidos utilizando um espectropolarímetro Jasco 710, operando a 25 °C. A concentração de proteína utilizada nas medidas de CD foram de 1,5 mg/mL em tampão (10 mM acetato de sódio, 20 mM NaCl e pH 5,5), utilizando uma cubeta de 0,1 mm de caminho óptico. Cada espectro representa uma media de 10 acumulações coletados na faixa de comprimentos de onda entre 190 a 260 nm, com um “step size” 0,2 nm, a uma velocidade de 20 nm/min, e uma largura de banda de 1 nm. A linha de base foi corrigida pela subtração do espectro do tampão

obtido em idênticas condições. Os resultados foram convertidos em elipticidade em unidades de mg/mL,  $[\Phi]$  (deg cm<sup>2</sup> decagrama<sup>-1</sup>), e a fração foi analisada com o programa CPro (SREERAMA e WOODY, 2000). Tendo em vista que os resultados fornecidos pelos métodos utilizados pelos programas SELCON3, CDSSTR, e CONTINL são similares, os valores apresentados relativos às frações de estrutura secundária da proteína refletem as médias dos resultados obtidos com os três programas.

#### **4.18 Espalhamento Dinâmico de Luz (Dynamic Light Scattering – DLS)**

A concentração de proteína utilizada nas medidas de DLS foi de 3.0 mg/mL em tampão contendo 30 mM de acetato de sódio, 30 mM de NaCl e 1% de glicerol, pH 5.5. As medidas de DLS foram realizadas utilizando-se o equipamento *SpectroLIGHT* 500 system, Molecular Dimensions, UK, portando um controlador de temperatura, em comprimento de onda de 781,2 nm e ângulo de detecção do espalhamento de 90°. O programa usado para a execução das medidas do raio hidrodinâmico foi o CONTIN. Antes do experimento, a amostra foi centrifugada a 15000xg por 10 min e a 10°C, a fim de remover possíveis partículas em suspensão.

#### **4.19 Teste de solubilidade da CP do CoLV em diferentes condições**

Uma amostra de proteína purificada CP-CoLV (3 mg/mL) foi utilizada para o teste de solubilidade. No ensaio, foram utilizados 24 diferentes tampões conforme o Quadro 2. Foi utilizada uma placa de cultura contendo 24 poços (reservatórios). No reservatório foi aplicado 1 mL de solução tampão e, na lamínula, adicionou-se 1 µL de proteína pura concentrada e 1 µL da solução do poço. Cada poço foi coberto com a respectiva lamínula de vidro, usando silicone para vedação.

<b>Condições</b>		
1. Glicina pH 3,0 0,1M	9. Fosfato de sódio pH 6,0 0,1M	17. EPPS pH 8,0 0,1M
2. Ácido cítrico pH 5,2 0,1M	10. Fosfato de sódio pH 7,0 0,1M	18. Imidazol pH 8,0 0,1M
3. Ácido cítrico pH 4,0 0,1M	11. MES pH 5,8 0,1M	19. Tris-HCl pH 8,5 0,1M
4. Acetato de sódio pH 4,0 0,1M	12. MES pH 6,2 0,1M	20. Bicina pH 8,5 0,1M
5. Acetato de sódio pH 5,0 0,1M	13. MES pH 6,5 0,1M	21. Bicina pH 9,0 0,1M
6. Citrato de sódio pH 4,7 0,1M	14. Cacodilato de sódio pH 6,5 0,1M	22. CAPS pH 10,0 0,1M
7. Citrato de sódio pH 5,5 0,1M	15. ADA pH 6,5 0,1M	23. HEPES pH 7,0 0,1M
8. Fosfato de sódio pH 5,0 0,1M	16. MOPS pH 7,0 0,1M	24. HEPES pH 7,5 0,1M

**Quadro 2.** Soluções utilizadas para o teste de solubilidade da CP-CoLV.

#### **4.20 Teste de cristalização das proteínas purificadas**

Inicialmente, Kits comerciais de soluções de cristalização foram utilizados na busca de condições iniciais de cristalização utilizando o método “hanging-drop” (gota suspensa). Os Kits utilizados foram: (1) Polietileno Glicol 6000 (Hampton Research); (2) Sulfato de Amônio (Hampton Research); PEG (Qiagen). Para os testes, foram utilizadas caixas plásticas para cultura de células contendo 24 reservatórios. 1 µl da solução de proteína foi misturado a igual volume da solução do Kit em uma lamínula. A lamínula foi colocada sobre um reservatório contendo 0,5 mL da mesma solução utilizando-se silicone para vedação.

#### **4.21 Gel filtração das frações eluídas para remoção do imidazol**

As frações eluídas foram aplicadas em coluna “Hiperp Desalting”

com matriz de gel filtração Sephadex G-25 (GE Healthcare) inicialmente equilibrada com o tampão desejado. Foram aplicados 10 mL de amostra à coluna e a eluição foi feita com tampão  $\text{NaH}_2\text{PO}_4$ , NaCl 50 mM, pH 7,4. O acompanhamento da cromatografia foi realizado fotometricamente em 280 nm.

#### **4.22 Predição de estrutura por ferramentas de bioinformática**

Utilizando a sequência de aminoácidos da proteína CP-CoLV (Anexo 2), foi obtido um modelo tridimensional por modelagem molecular comparativa. Para as regiões com similaridade >30% de sequência os modelos foram montados pela modelagem por homologia e para as regiões com similaridade <30%, os modelos foram montados utilizando a técnica de Threader (JONES, 1999; MCGUFFIN e JONES, 2003; LOBLEY et al., 2009), disponível no servidor do PSIPRED (BUCHAN et al., 2013). Os modelos foram construídos por alinhamento com a estrutura-molde no programa Modeller (ESWAR et al., 2006). Os modelos foram submetidos à avaliação da geometria utilizando o gráfico de Ramachandran (RAMACHANDRAN et al., 1963). Para a avaliação da qualidade do modelo quanto a estereoquímica, foi utilizado o programa WHAT\_CHECK (HOOFT et al., 1996) e PROCHECK (LASKOWSKI et al., 1993).

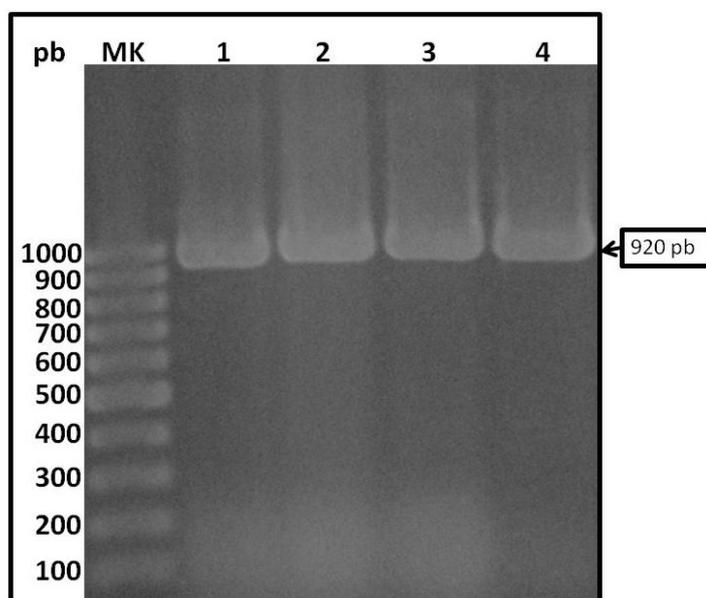
A predição de estrutura secundária, análises estruturais (acessibilidade, pontes), modificações pós-traducionais, dentre outras foram analisadas no pacote PredicProtein (ROST et al., 2004).

## 5. Resultados e Discussão

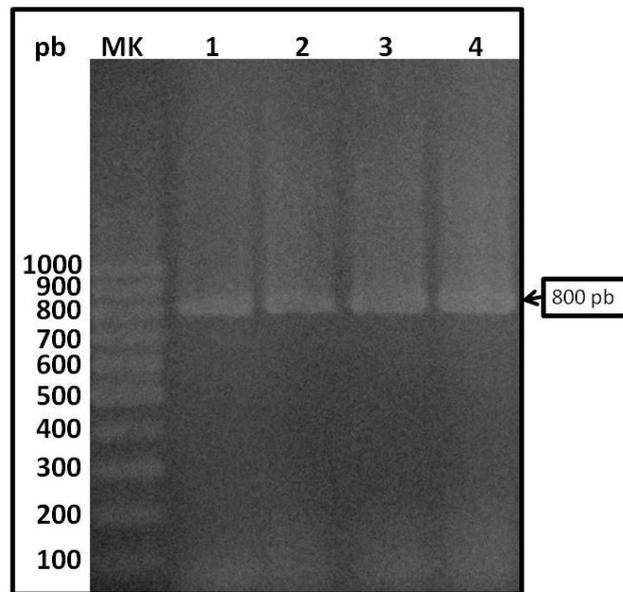
### 5.1 Transformação dos plasmídeos recombinantes em células competentes

#### 5.1.1 Análise das colônias por PCR

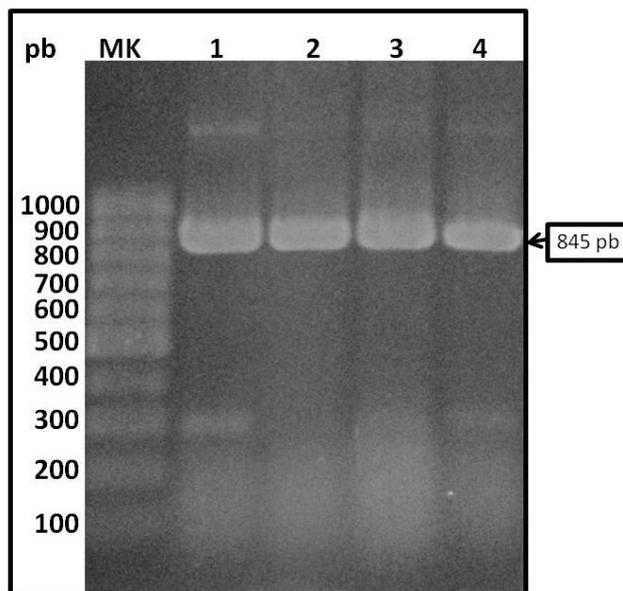
Após cada transformação, quatro colônias foram escolhidas para os testes confirmatórios por PCR do fragmento alvo. As figuras 9-11 são fotografias do gel de eletroforese em agarose dos produtos amplificados. A figura 9 mostra os produtos de 920 pb referentes ao gene que codifica a CP do CoLV; a figura 10 mostra os produtos de 800 pb que correspondem ao gene que codifica a MP do PepRSV e a figura 11 refere-se aos produtos de 845 pb que codificam a helicase do PepRSV. Todos os produtos tiveram tamanho condizente com o esperado.



**Figura 9.** Gel nativo de agarose 1% evidenciando em: (1, 2, 3, 4) fragmento de 920 pb amplificado por PCR referente ao gene da CP do CoLV;; (MK) Marcador molecular.



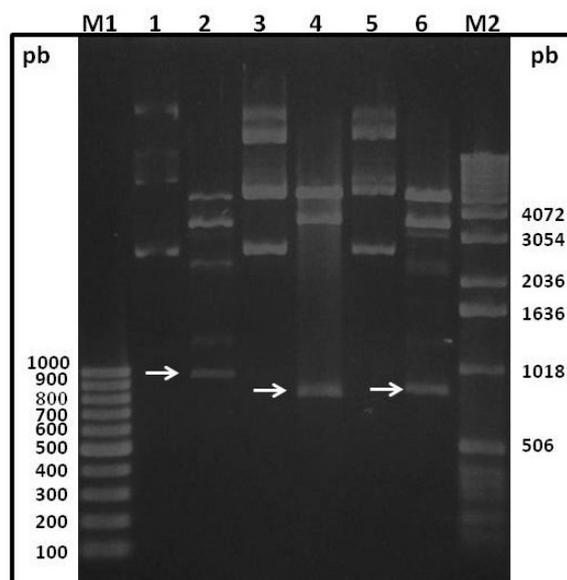
**Figura 10.** Gel nativo de agarose 1% evidenciando em: (1, 2, 3, 4) fragmento de 800 pb amplificado por PCR referente ao gene da MP do PepRSV (MK) Marcador molecular.



**Figura 11.** Gel nativo de agarose 1% evidenciando em: (1, 2, 3, 4) fragmento de 845 pb amplificado por PCR referente ao gene da helicase do PepRSV; (MK) Marcador molecular GeneRuler 100pb DNA Ladder (Fermentas).

### 5.1.2 Análise dos plasmídeos por restrição enzimática

O DNA plasmidial, de um dos clones positivos para a reação de PCR, foi extraído e analisado por restrição enzimática para uma segunda confirmação da presença do inserto de interesse. Os três pares de oligonucleotídeos específicos apresentam na orientação senso o sítio de restrição para a enzima Bam HI e na orientação antissenso o sítio de restrição para a enzima Hind III. A figura 12 apresenta o perfil de restrição dos três clones selecionados em gel de agarose. As colônias de bactérias que, tiveram seus DNAs plasmidiais extraídos para a realização deste teste, foram utilizadas para os testes de expressão.



**Figura 12.** Gel nativo de agarose 1% evidenciando em: (1) vetor pET 28a ligado ao gene da CP do CoLV, antes da digestão enzimática; (2) digestão enzimática do vetor com liberação de fragmento de 920 pb; (3) vetor pET 28a ligado ao gene da MP do PepRSV, antes da digestão enzimática; (4) digestão enzimática do vetor com liberação de fragmento de 800 pb; (5) vetor pET 28a ligado ao gene da HEL do PepRSV, antes da digestão enzimática; (6) digestão enzimática do vetor com liberação de fragmento de 845 pb; (M1) Marcador molecular de 100pb; (M2) Marcador molecular de 1 Kb. As setas indicam os fragmentos liberados.

### 5.2 Testes de expressão e solubilidade

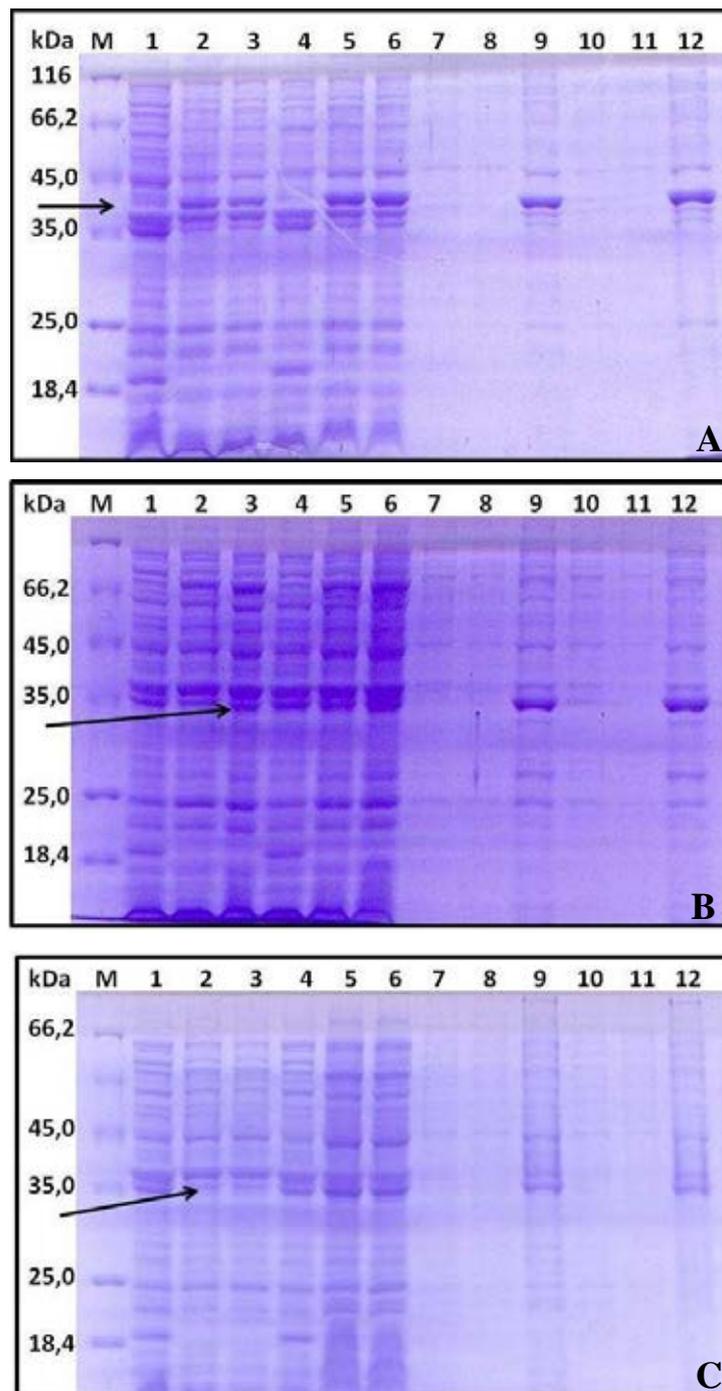
O estudo estrutural das proteínas virais depende da obtenção de altas concentrações dessas proteínas e elevado grau de pureza. Em geral, as

proteínas de fitovírus estão presentes nas plantas em baixas concentrações e podem não ser estáveis dependendo da função que exercem. Para driblar a dificuldade em obter tais proteínas através de extratos vegetais, nesse trabalho, optou-se por expressá-las utilizando células de *E. coli*.

*E. coli* é, ainda, a principal hospedeira no que diz respeito a expressão de proteínas heterólogas. Isso, devido à simplicidade e velocidade para a produção das proteínas. Como desvantagem, podemos citar as grandes chances de formação de corpos de inclusão (agregado de proteínas insolúveis) (MIDDELBERG, 2002). Efetivamente, apenas de 1 a 5% das ORFs expressadas em sistemas heterólogos tiveram suas estruturas resolvidas. Isso se deve, principalmente, à dificuldade de obtenção de proteínas solúveis (SHEN, et al., 2005).

A expressão de proteínas solúveis é uma grande vantagem, uma vez que, essas proteínas apresentam-se em seu estado nativo. Porém, na maioria dos casos, a maior parte das proteínas produzidas agrega-se formando os corpos de inclusão. Os principais fatores responsáveis pela formação dos agregados são: interações hidrofóbicas, pontes dissulfeto e altas taxas de expressão. Tanto as interações hidrofóbicas quanto os altos níveis de produção de proteínas são favorecidos quando a expressão é feita a 37°C. Sendo assim, é possível, em alguns casos, aumentar a solubilidade da proteína de interesse reduzindo a temperatura de cultivo, uma vez que a taxa metabólica das bactérias é reduzida (SCHEIN e NOTEBOM, 1988).

Neste sentido, foram realizados testes de expressão em diferentes temperaturas e diferentes concentrações de IPTG, a fim de, avaliar a solubilidade das proteínas e as taxas de expressão. Quando a expressão foi realizada a 37°C, observou-se altos níveis de expressão para as três proteínas estudadas, no entanto, em sua totalidade, as proteínas formaram corpos de inclusão. Também observamos que, as taxas de expressão não foram significativamente alteradas com a variação na concentração de IPTG. A figura 13 ilustra os testes de expressão realizados, respectivamente, para: a CP do CoLV (Figura 13A); MP do PepRSV (Figura 13B); HEL do PepRSV (Figura 13C).

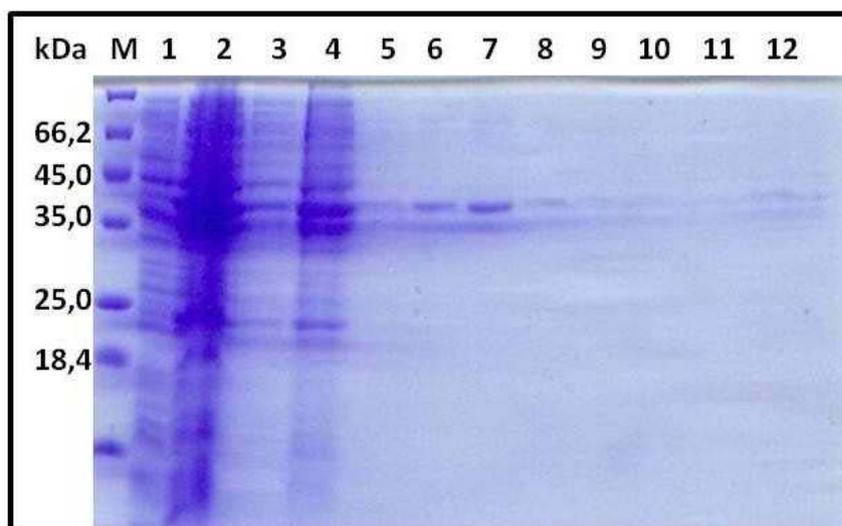


**Figura 13. (A)** SDS-PAGE 12% mostrando o teste de expressão a 37°C para a proteína capsial do CoLV; **(B)** SDS-PAGE 12% mostrando o teste de expressão a 37°C para a proteína do movimento do PepRSV; **(C)** SDS-PAGE 12% mostrando o teste de expressão a 37°C para a helicase do PepRSV (M) Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas); (1 e 4) Extratos protéicos totais não induzidos no sistema de expressão; (2) Extrato total após 4h da indução com 0,5 mM de IPTG; (3) Extrato total após 4h da indução com 1,0 mM de IPTG; (5) Extrato total após 18h da indução com 0,5 mM de IPTG; (6) Extrato total após 18h da indução com 1,0 mM de IPTG; (7 e 8) Frações solúveis após, respectivamente, 1ª e 2ª lises celulares da amostra induzida com 0,5 mM de IPTG; (9) Fração insolúvel, pós lise celular, da amostra induzida com 0,5 mM de IPTG; (10 e 11) Frações solúveis após, respectivamente, 1ª e 2ª lises celulares da amostra induzida com 1,0 mM de IPTG; (12) Fração insolúvel, pós lise celular, da amostra induzida com 1,0 mM de IPTG.

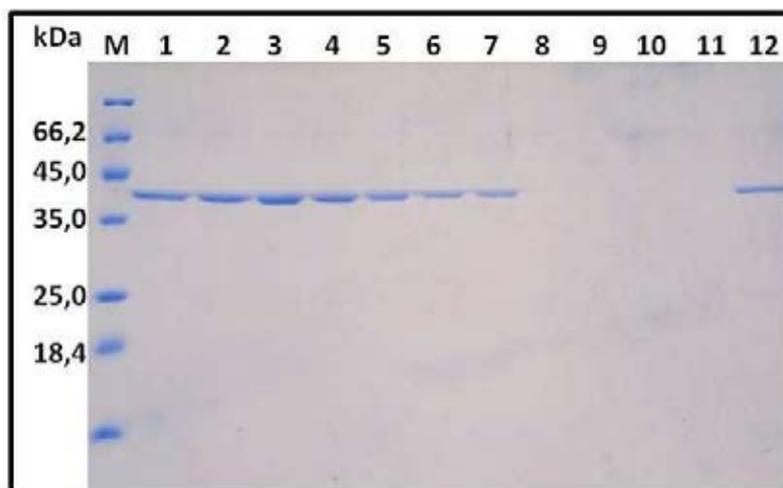
Os testes de expressão realizados a 25°C mostraram que a esta temperatura é possível obter CP do CoLV na fração solúvel, pós lise celular, porém, em quantidades insignificantes quando comparadas a presente na fração insolúvel. Tanto a MP quanto a Helicase do PepRSV permaneceram apenas na fração insolúvel. Para todas as proteínas foi possível observar uma queda nos níveis de expressão conforme já era esperado (dados não apresentados).

Após a observação de provável solubilização da CP do CoLV nos testes de expressão realizados a 25°C, iniciaram-se novos testes de expressão reduzindo a temperatura para 18°C. Para confirmar a presença/ausência de proteínas nas frações solúveis e insolúveis, essas amostras foram purificadas por cromatografia em coluna de afinidade contendo resina de níquel. As frações solúveis foram purificadas em condições não desnaturantes e as frações insolúveis em condições desnaturantes.

Os experimentos realizados com a CP do CoLV a 18°C mostraram que é possível purificar essa proteína tanto em condições nativas (Figura 14) quanto em condições desnaturantes (Figura 15). Entretanto, fica evidente que a concentração de proteínas purificadas a partir da fração insolúvel ainda é maior que a obtida a partir da fração solúvel.

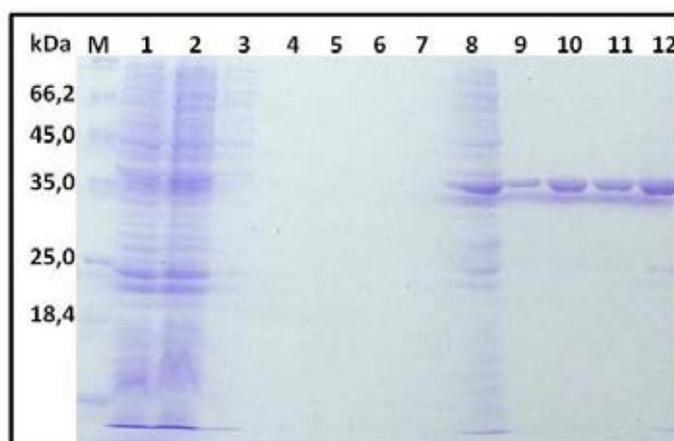


**Figura 14.** SDS-PAGE 12% mostrando os resultados dos testes de expressão a 18°C e purificação da CP do CoLV em condições nativas. (M) Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas); (1) Extratos protéicos totais não induzidos no sistema de expressão; (2) Extrato total após 18h da indução com 1,0 mM de IPTG; (3) Fração solúvel, após lise celular; (4) Fração insolúvel, após lise celular; (5-12) Eluições resultantes da purificação da fração solúvel.



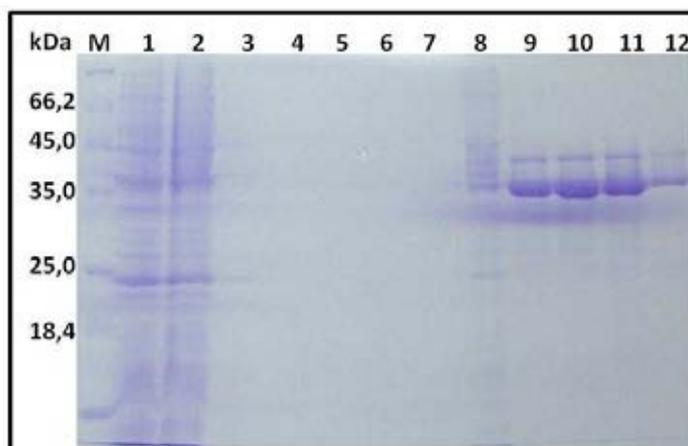
**Figura 15.** SDS-PAGE 12% mostrando o resultado da purificação da CP do CoLV em condições desnaturantes. (M) Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas); (1-11) Eluições resultantes da purificação da fração insolúvel; (12) Amostra da resina retirada após as eluições.

Os resultados dos testes realizados a 18°C com a MP e a HEL do PepRSV confirmaram que, mesmo em expressão a baixas temperaturas, essas proteínas permanecem insolúveis. A figura 16 evidencia que não foi possível obter MP a partir da fração solúvel purificada em condições nativas, no entanto, obteve-se quantidades razoáveis e em bom grau de pureza nas eluições da purificação desnaturante.



**Figura 16.** SDS-PAGE 12% mostrando os resultados do teste de expressão a 18°C e purificação da MP do PepRSV. (M) Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas); (1) Extratos protéicos totais não induzidos no sistema de expressão; (2) Extrato total após 18h da indução com 1,0 mM de IPTG; (3) Fração solúvel, após lise celular; (4-7) Eluições resultantes da purificação da fração solúvel; (8) Fração insolúvel, após lise celular; (9-12) Eluições resultantes da purificação da fração insolúvel.

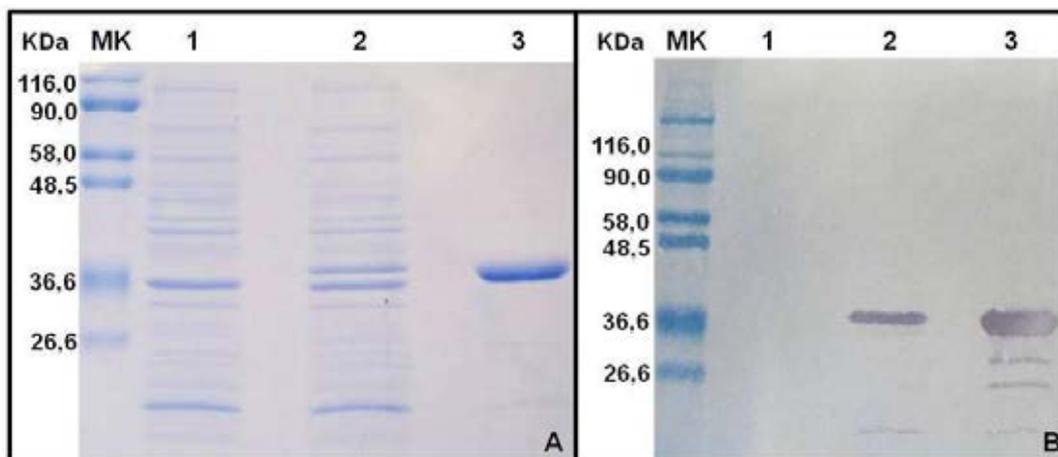
Assim como para a MP, também não foi possível obter a Helicase do PepRSV em expressão a baixas temperaturas. A proteína de interesse foi purificada em condições desnaturantes, conforme mostra a Figura 17. Fica evidente a presença de uma banda não específica acima da banda de interesse. Outros testes já foram realizados, utilizando tampões de lavagem com baixos valores de pH, o que favoreceu a eliminação desta banda contaminante.



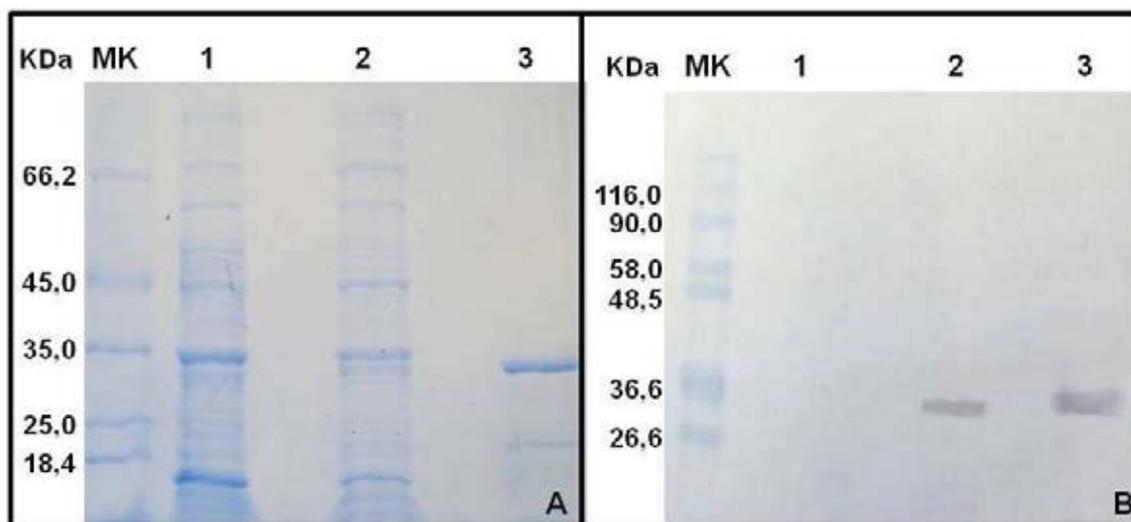
**Figura 17.** SDS-PAGE 12% mostrando os resultados do teste de expressão a 18°C e purificação da Helicase do PepRSV. (M) Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas); (1) Extratos protéicos totais não induzidos no sistema de expressão; (2) Extrato total após 18h da indução com 1,0 mM de IPTG; (3) Fração solúvel, após lise celular; (4-7) Eluições resultantes da purificação da fração solúvel; (8) Fração insolúvel, após lise celular; (9-12) Eluições resultantes da purificação da fração insolúvel.

### 5.3 Confirmação da purificação das proteínas de interesse por western-blot

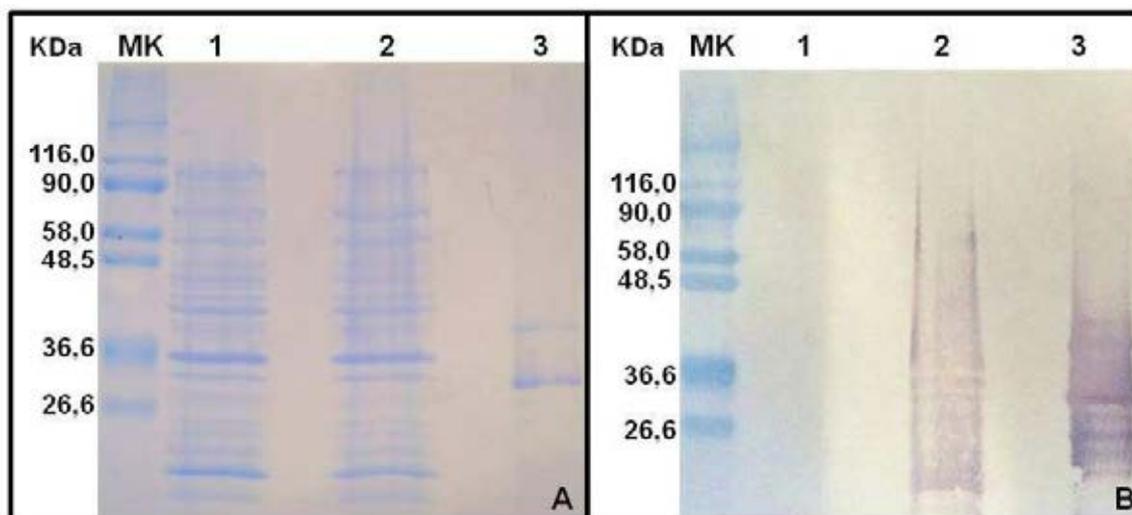
Um gel idêntico ao das Figuras 18A, 19A e 20A foi transferido para membrana de nitrocelulose. A verificação da expressão das proteínas virais de interesse foi realizada através do western-blot de utilizando o antissoro anti-poli-histidina (Figuras 19B, 20B e 21B), onde a reação foi específica conforme o esperado com as amostras do extrato de células com 4 horas de indução e com as amostras das proteínas recombinantes purificadas.



**Figura 18 (A)** SDS-PAGE 12% mostrando: (M) Marcador de massa molecular (kDa) PageRuler Prestained Protein Ladder (Fermentas); (1) Extrato protéico total não induzidos no sistema de expressão da CP do CoLV. (2) Extrato proteico total após 4h da indução com 1,0 mM de IPTG; (3) Proteína CP-CoLV purificada em condições desnaturantes. **(B)** Western blot de um gel idêntico ao da Figura 18 A evidenciando a expressão das proteínas de interesse.



**Figura 19 (A)** SDS-PAGE 12% mostrando: (M) Marcador de massa molecular (kDa) PageRuler Prestained Protein Ladder (Fermentas); (1) Extrato protéico total não induzido no sistema de expressão referente ao teste da MP do PepRSV; (2) Extrato total após 4h da indução com 1,0 mM de IPTG; (3) Proteína purificada em condições desnaturantes. **(B)** Western blot de um gel idêntico ao da Figura 19 A evidenciando a expressão da proteína de interesse.



**Figura 20 (A)** SDS-PAGE 12% mostrando: (M) Marcador de massa molecular (kDa) PageRuler Prestained Protein Ladder (Fermentas); (1) Extrato protéico não induzido no sistema de expressão referente ao teste da Hel do PepRSV; (2) Extrato total após 4h da indução com 1,0 mM de IPTG; (3) Proteína purificada em condições desnaturantes. **(B)** Western blot de um gel idêntico ao da Figura 20 A evidenciando a expressão da proteína de interesse.

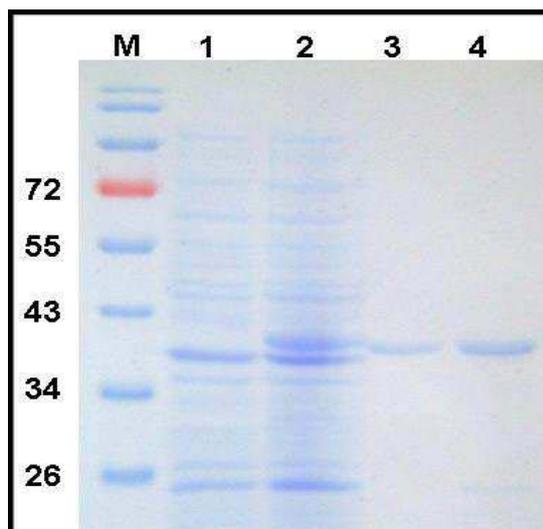
#### 5.4 Reenovelamento das proteínas por gradiente 8-0 M uréia

Uma das estratégias utilizadas para a obtenção de proteínas enoveladas a partir de corpos de inclusão é a solubilização destes agregados em um agente caotrópico (como uréia ou guanidina) em altas concentrações (6 – 8 M) e ligação das proteínas recombinantes em matriz de afinidade, seguidas pelo *refolding* através de lavagens com tampões em gradiente decrescente do agente caotrópico (COLANGELI et al., 1998). A eluição das proteínas é feita em tampão não desnaturante.

Foram realizados experimentos de reenovelamento por gradiente para as três proteínas em estudo. Após a tentativa de *refolding* em coluna, tanto a MP quanto a Helicase do PepRSV permaneceram ligadas à resina após a etapa de eluição. Uma possível explicação para esse fato seria a ineficiência do processo de reenovelamento, fazendo com que as proteínas permanecessem em estado desnaturado ou semi-desnaturado e, dessa forma, o tampão de eluição para proteínas nativas não seria capaz de desligá-las da matriz de níquel.

Quando o mesmo experimento foi realizado para purificação e

reenovelamento da CP do CoLV, foi possível obter proteínas em bom estado de pureza e supostamente renaturadas. A figura 21 mostra uma das eluições obtida neste experimento em comparação com uma eluição obtida por purificação em condições nativas. Para verificar a presença de estruturas secundárias nas proteínas eluídas, ambas as amostras foram analisadas por espectroscopia de infravermelho.



**Figura 21:** Gel de poliacrilamida (SDS-PAGE) 12% mostrando a expressão em *E. coli* e purificação em resina de níquel da proteína capsidial do CoLV fusionada a cauda de histidina. (M) Marcador de massa molecular (kDa) PageRuler Prestained Protein Ladder (Fermentas) (1) Amostra de células antes da indução da expressão. (2) Amostra de células após 4 horas de indução com 1mM de IPTG a 18°C. (3) Proteína purificada em condições nativas. (4) Proteína purificada em condições desnaturantes com posterior reenovelamento.

### 5.5 Espectroscopia de infravermelho (FT-IR) da CP do CoLV purificada em condições nativas e em condições desnaturantes com posterior reenovelamento

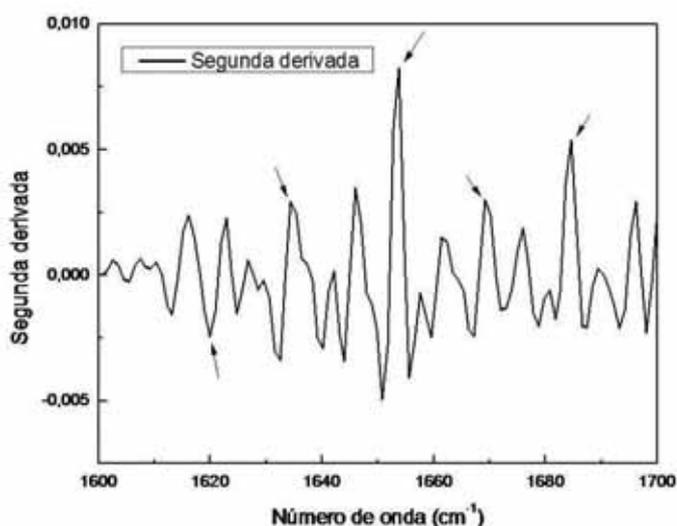
A espectroscopia de infravermelho é uma técnica que, nas últimas décadas, tem sido amplamente reconhecida para a investigação de estruturas secundárias, enovelamento e estabilidade de proteínas (ARRONDO e GONI, 1999; DONG et al., 1990).

As estruturas secundárias das proteínas podem ser estudadas usando o espectro de absorção da proteína na região da banda de amida I (1700-1600  $\text{cm}^{-1}$ ) que, corresponde à absorção das ligações peptídicas C=O (DOGLIA

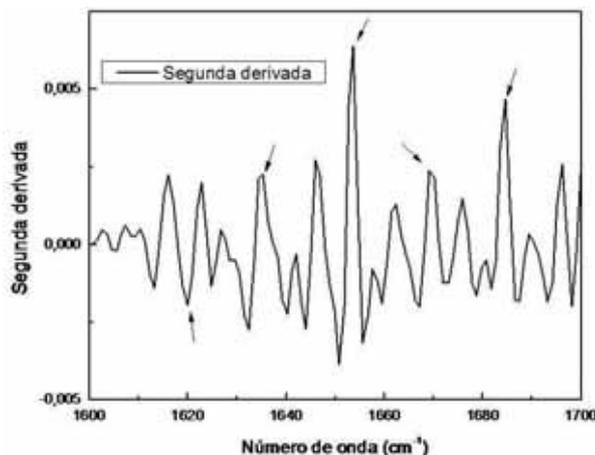
et al., 2008). Através deste espectro de absorção é possível verificar a porcentagem de cada tipo de estrutura secundária presente na amostra (Folha- $\beta$ , Hélice- $\alpha$ , Folha- $\beta$  anti-paralela, Voltas- $\beta$  e agregados).

As análises por espectroscopia de infravermelho da CP recombinante do CoLV purificada em condições nativas e da CP recombinante do CoLV purificada do corpo de inclusão com posterior reenovelamento permitiram comparar o estado conformacional das proteínas obtidas pelos dois processos. Os resultados podem ser observados através do gráfico da segunda derivada conforme mostrado na figura 22 e 23 e as porcentagens de estruturas secundárias a partir do gráfico da deconvolução para ambas proteínas como mostrado na figura 24 e 25.

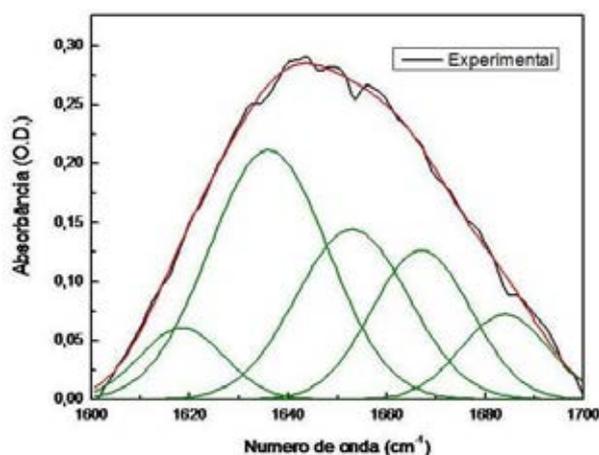
As análises por infravermelho indicaram a presença de estruturas secundárias constituídas de: 41% de folhas- $\beta$ , 22% de hélices- $\alpha$ , 20% de folhas- $\beta$  anti-paralelas, 9% de voltas- $\beta$  e 7% de agregados para a proteína purificada em condições nativas (Tabela 3) e 38% de folhas- $\beta$ , 24% de hélices- $\alpha$ , 19% de folhas- $\beta$  anti-paralelas, 10% de voltas- $\beta$  e 8% de agregados para a proteína purificada de forma desnaturante com posterior reenovelamento (Tabela 4). Esses resultados mostraram, ainda, um alto grau de similaridade conformacional entre as proteínas obtidas nos dois processos de purificação.



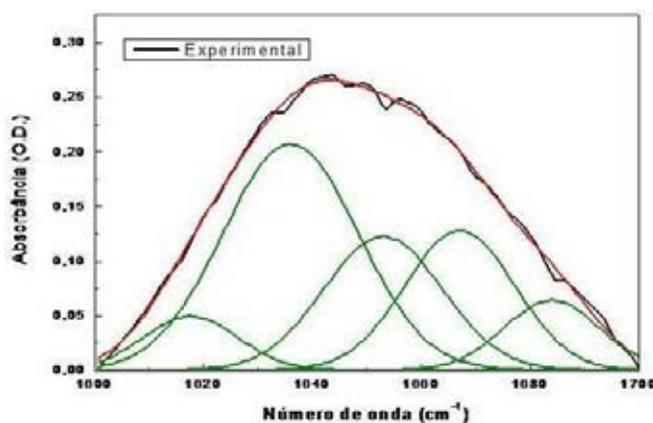
**Figura 22.** Gráfico da segunda derivada da proteína capsial do CoLV quando purificada de forma nativa. As setas indicam as posições das contribuições mais significativas do modo vibracional C=O na proteína.



**Figura 23.** Gráfico da segunda derivada da proteína capsial do CoLV quando purificada de forma desnaturante com posterior reenovelamento. As setas indicam as posições das contribuições mais significativas do modo vibracional C=O na proteína.



**Figura 24.** Gráfico da deconvolução espectral da banda de Amida I mostrando as contribuições indicadas pelo gráfico de segunda derivada da proteína do CoLV (Figura 11). A área abaixo de cada curva representa o percentual da contribuição de cada estrutura secundária.



**Figura 25.** Gráfico da deconvolução espectral da banda de Amida I mostrando as contribuições indicadas pelo gráfico de segunda derivada da proteína do CoLV (Figura 12). A área abaixo de cada curva representa o percentual da contribuição de cada estrutura secundária.

**Tabela 3.** Cálculo das áreas e respectivas porcentagens das bandas gaussianas obtidas por deconvolução espectral da proteína capsial do CoLV purificada em condições nativas.

Estrutura secundária	Posição (cm <sup>-1</sup> )	Percentual
Agregados	1618	7
Folha-β	1635	41
Hélice-α	1653	22
Folha-β anti-paralela	1667	20
Voltas-β	1684	9

**Tabela 4.** Cálculo das áreas e respectivas porcentagens das bandas gaussianas obtidas por deconvolução espectral da proteína capsial do CoLV purificada em condiçõesdesnaturantes com posterior reenovelamento.

Estrutura secundária	Posição (cm <sup>-1</sup> )	Percentual
Agregados	1618	8
Folha-β	1636	38
Hélice-α	1653	24
Folha-β anti-paralela	1667	19
Voltas-β	1684	10

## 5.6 Testes de solubilidade com agentes não-desnaturantes

Uma vez que, as estratégias de reenovelamento por gradiente, diluição ou diálise não foram eficientes para a MP e Helicase do PepRSV, outras tentativas de solubilização dessas proteínas a partir dos corpos de inclusão foram feitas.

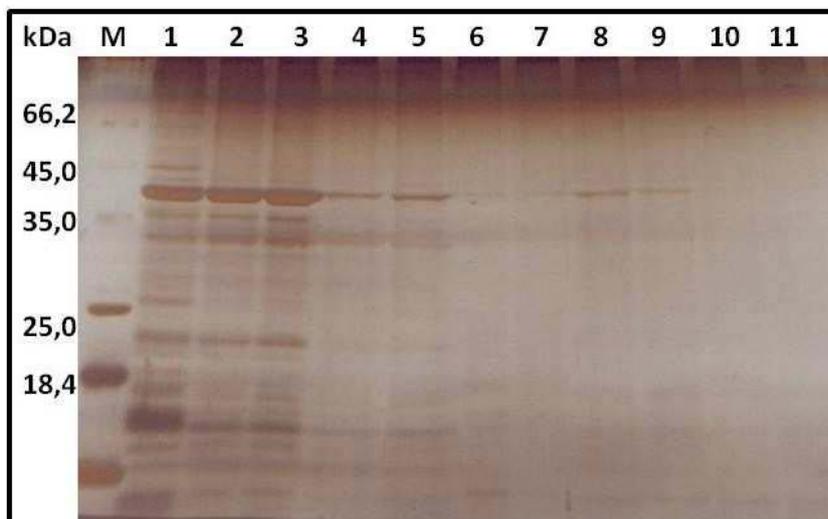
Foi observado por TSUMOTO et al. (2004) que a adição de 0,1-1,0 M de arginina no tampão de lise aumenta a quantidade de proteínas na fração solúvel. Nos testes de solubilização realizados, além da argina, foram utilizados outros três tipos tipos de agentes não-desnaturantes (ciclodextrina, sulfobetaina e PEG).

A fim de verificar a capacidade de solubilização apenas da fração insolúvel, foi feita uma lise celular inicial sem a adição dos reagentes em questão e, após a remoção da primeira fração solúvel, os tampões contendo os agentes

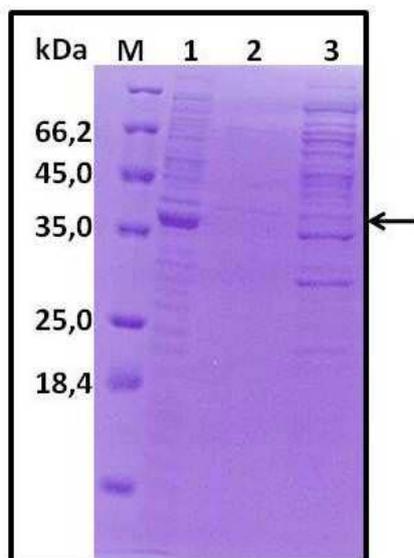
não-desnaturantes foram utilizados.

Como resultados, podemos observar na Figura 26 que, a arginina (0,5-1,0 M) foi o reagente mais eficiente na solubilização da CP do CoLV. Os agentes ciclodextrina e sulfobetaina também proporcionaram, em menor grau, a obtenção de proteínas solúveis. Fica evidente que a arginina solubiliza, praticamente, apenas a proteína de interesse e poucas proteínas não específicas. Essa estratégia é de grande interesse, pois dessa forma, seria possível a obtenção da proteína solúvel purificada, sem a etapa de cromatografia de afinidade.

Levando-se em consideração o fato de que os corpos de inclusão obtidos a baixas temperaturas podem apresentar grande parte das proteínas em estado similar ao nativo, foi feito um novo teste de solubilização para a CP do CoLV com corpos de inclusão obtidos por expressão a 18°C. A figura 27 ilustra o resultado obtido, verificando-se que a fração solubilizada é composta por inúmeros tipos de proteínas não específicas.

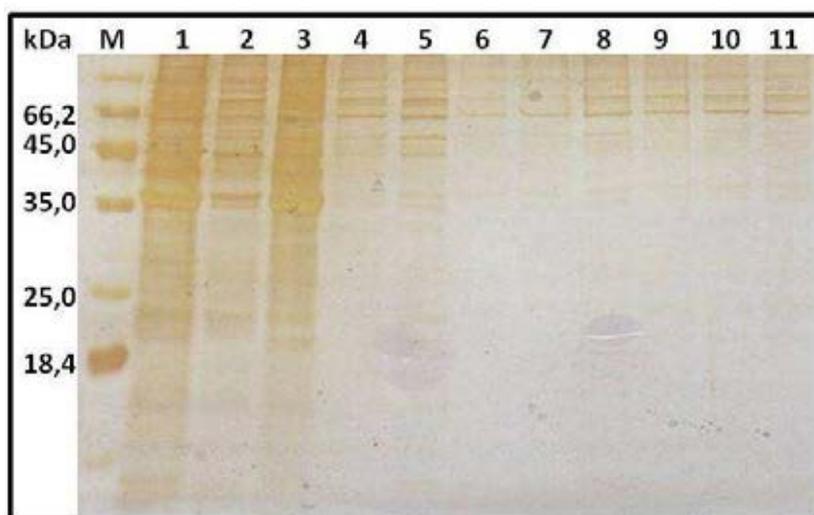


**Figura 26.** SDS-PAGE 12% mostrando as tentativas de solubilização dos corpos de inclusão (CP CoLV) obtidos a 18°C, com diferentes agentes não desnaturantes. **(M)** Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas). **(1)** Fração insolúvel após lise celular; **(2)** Fração insolúvel após lavagens em tampão contendo 2% deoxicolato; **(3)** Fração insolúvel após lavagens em tampão contendo 1% triton X-100; **(4)** Solubilização com Arginina 0,5M; **(5)** Solubilização com Arginina 1,0 M; **(6)** Sulfobetaina 0,5M; **(7)** Sulfobetaina 1,0 M; **(8)** Ciclodextrina 50 mM; **(9)** Ciclodextrina 100mM; **(10)** PEG 3300 0,1g/L; **(11)** PEG 3300 0,5g/L.

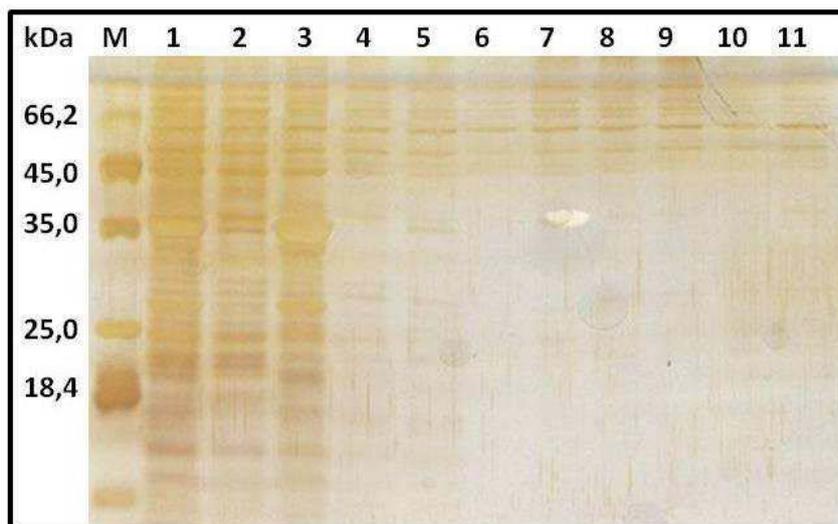


**Figura 27.** SDS-PAGE 12% mostrando as tentativas de solubilização na presença de arginina, dos corpos de inclusão (CP CoLV) obtidos a 18°C. **(M)** Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas); **(1)** Fração insolúvel após lise celular e lavagens; **(2)** Fração solubilizada a partir de corpos de inclusão obtidos em expressão a 37°C com arginina 1,0 M; **(3)** Fração solubilizada a partir de corpos de inclusão obtidos em expressão a 18°C com arginina 1,0M.

Não foram obtidos resultados satisfatórios de solubilização para a MP (Figura 28) e para a Helicase (Figura 29) do PepRSV.



**Figura 28.** SDS-PAGE 12% mostrando as tentativas de solubilização dos corpos de inclusão (MP do PepRSV) obtidos a 37°C, com diferentes agentes não desnaturantes. **(M)** Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas). **(1)** Fração insolúvel após lise celular; **(2)** Sobrenadante da primeira lavagem em tampão contendo 2% deoxicolato; **(3)** Fração insolúvel após lavagens em tampão contendo 1% triton X-100; **(4)** Solubilização com Arginina 0,5M; **(5)** Solubilização com Arginina 1,0 M; **(6)** Sulfobetaina 0,5M; **(7)** Sulfobetaina 1,0 M; **(8)** Ciclodextrina 50 mM; **(9)** Ciclodextrina 100mM; **(10)** PEG 3300 0,1g/L; **(11)** PEG 3300 0,5g/L.



**Figura 29.** SDS-PAGE 12% mostrando as tentativas de solubilização dos corpos de inclusão (HEL do PepRSV) obtidos a 37°C, com diferentes agentes não desnaturantes. **(M)** Marcador para proteínas, Unstained Protein Molecular Weight Marker (Fermentas). **(1)** Fração insolúvel após lise celular; **(2)** Sobrenadante da primeira lavagem em tampão contendo 2% deoxicolato; **(3)** Fração insolúvel após lavagens em tampão contendo 1% triton X-100; **(4)** Solubilização com Arginina 0,5M; **(5)** Solubilização com Arginina 1,0 M; **(6)** Sulfobetaina 0,5M; **(7)** Sulfobetaina 1,0 M; **(8)** Ciclodextrina 50 mM; **(9)** Ciclodextrina 100mM; **(10)** PEG 3300 0,1g/L; **(11)** PEG 3300 0,5g/L.

### 5.7 Redução de proteínas não específicas através da lavagem dos corpos de inclusão

Os corpos de inclusão podem conter, além das proteínas, outros tipos de polipeptídeos, fosfolípidios e ácidos nucleicos, dependendo do sistema de expressão e condições de crescimento (VALAX e GEORGIU, 1993). A presença desses contaminantes pode reduzir e prejudicar, significativamente, o processo de renovelamento das proteínas de interesse (MAACHUPALLI-REDDY, et al., 1997).

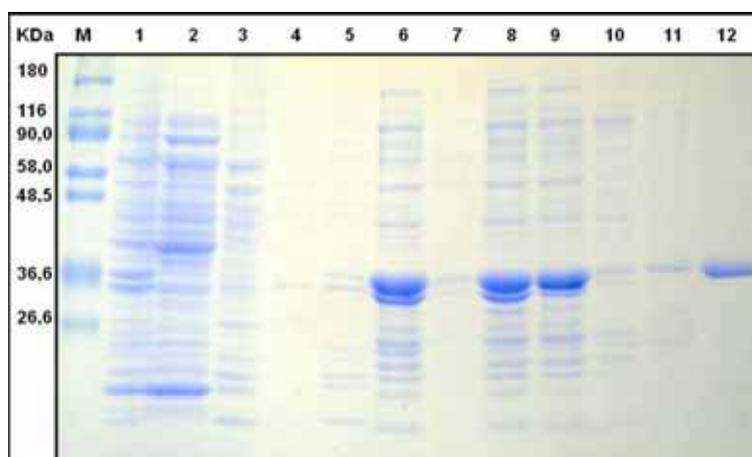
Foi observado que o processo mecânico/químico de lavagem dos corpos de inclusão com detergentes, uréia (0-4 M) ou sacarose contribui para diminuir os níveis de contaminantes nos corpos de inclusão (FISHER, et al., 1993).

Dessa forma, para as três proteínas, foram realizadas etapas de lavagem utilizando Triton X-100 a 1%, deoxicolato a 2% e uréia 0,5 M. Esse

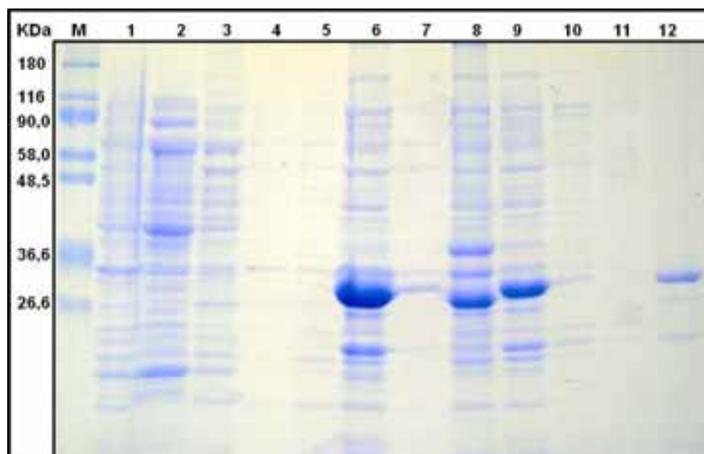
procedimento adicional contribuiu significativamente para o sucesso nos experimentos de reenovelamento das proteínas. Os géis a seguir (Figuras 30, 31 e 32, canaletas de 1 a 7), evidenciam as proteínas contaminantes eliminadas em cada etapa de lavagem. Provavelmente, outros contaminantes, como os ácidos nucleicos, também sejam eliminados durante as lavagens.

A lavagem dos corpos de inclusão contribuiu para a obtenção de amostras com menor concentração de proteínas contaminantes quando comparadas com os resultados apresentados no relatório anterior.

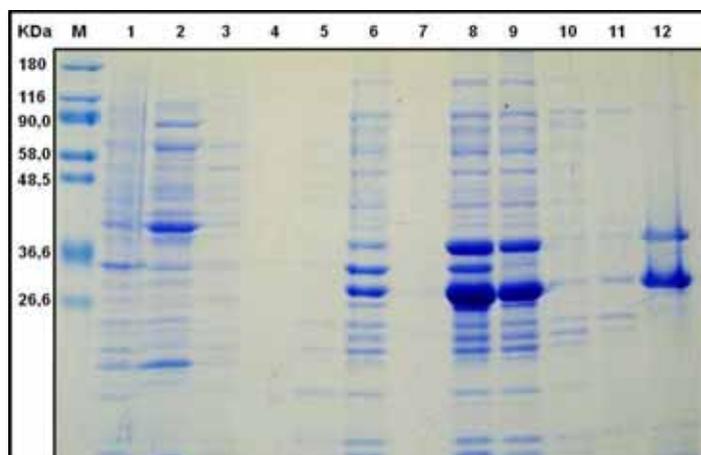
Os melhores resultados foram alcançados para a proteína CP-CoLV. A canaleta 12 da Figura 31 evidencia a amostra purificada livre de contaminantes. Para a proteína MP do PepRSV houve uma melhora importante para o grau de pureza mas ainda é possível observar a presença de baixa concentração de proteínas contaminantes (canaleta 12 da Figura 31). Para a purificação da HEL do PepRSV encontramos maior dificuldade, mesmo com as etapas de lavagem, ainda é evidente uma forte banda de contaminação como mostra a canaleta 12 da Figura 32.



**Figura 30.** SDS-PAGE 12% mostrando as etapas de lavagem do corpo de inclusão e as etapas de purificação da CP do CoLV. (M) Marcador para proteínas, Prestained Molecular Weight Marker (Sigma); (1) Suspensão dos corpos de inclusão em tampão deoxicolato 2% antes da lise celular; (2) e (4) Sobrenadantes a serem descartados após lavagem com tampão deoxicolato; (3) e (5) Sobrenadantes a serem descartados após lavagem com tampão Triton X-100; (6) Suspensão dos corpos de inclusão em tampão contendo 0,5M uréia; (7) Sobrenadante a ser descartado após lavagem com tampão 0,5M uréia; (8) Amostra resultante da solubilização dos corpos de inclusão em 8M uréia; (9) Solubilizado após centrifugação a 14.000 g a ser purificado; (10) Tampão coletado durante a etapa de lavagem da purificação em pH 6,3; (11) Tampão coletado durante a etapa de lavagem da purificação em pH 5,9; (12) Fração resultante da eluição da proteína.



**Figura 31.** SDS-PAGE 12% mostrando as etapas de lavagem do corpo de inclusão e as etapas de purificação da MP do PepRSV. (M) Marcador para proteínas, Prestained Molecular Weight Marker (Sigma); (1) Suspensão dos corpos de inclusão em tampão deoxicolato 2% antes da lise celular; (2) e (4) Sobrenadantes a serem descartados após lavagem com tampão deoxicolato; (3) e (5) Sobrenadantes a serem descartados após lavagem com tampão Triton X-100; (6) Suspensão dos corpos de inclusão em tampão contendo 0,5M uréia; (7) Sobrenadante a ser descartado após lavagem com tampão 0,5M uréia; (8) Amostra resultante da solubilização dos corpos de inclusão em 8M uréia; (9) Solubilizado após centrifugação a 14.000 g a ser purificado; (10) Tampão coletado durante a etapa de lavagem da purificação em pH 6,3; (11) Tampão coletado durante a etapa de lavagem da purificação em pH 5,9; (12) Fração resultante da eluição da proteína.



**Figura 32.** SDS-PAGE 12% mostrando as etapas de lavagem do corpo de inclusão e as etapas de purificação da HEL do PepRSV. (M) Marcador para proteínas, Prestained Molecular Weight Marker (Sigma); (1) Suspensão dos corpos de inclusão em tampão deoxicolato 2% antes da lise celular; (2) e (4) Sobrenadantes a serem descartados após lavagem com tampão deoxicolato; (3) e (5) Sobrenadantes a serem descartados após lavagem com tampão Triton X-100; (6) Suspensão dos corpos de inclusão em tampão contendo 0,5M uréia; (7) Sobrenadante a ser descartado após lavagem com tampão 0,5M uréia; (8) Amostra resultante da solubilização dos corpos de inclusão em 8M uréia; (9) Solubilizado após centrifugação a 14.000 g a ser purificado; (10) Tampão coletado durante a etapa de lavagem da purificação em pH 6,3; (11) Tampão coletado durante a etapa de lavagem da purificação em pH 5,9; (12) Fração resultante da eluição da proteína.

## 5.8 Renovelamento das proteínas e análise por dicroísmo circular

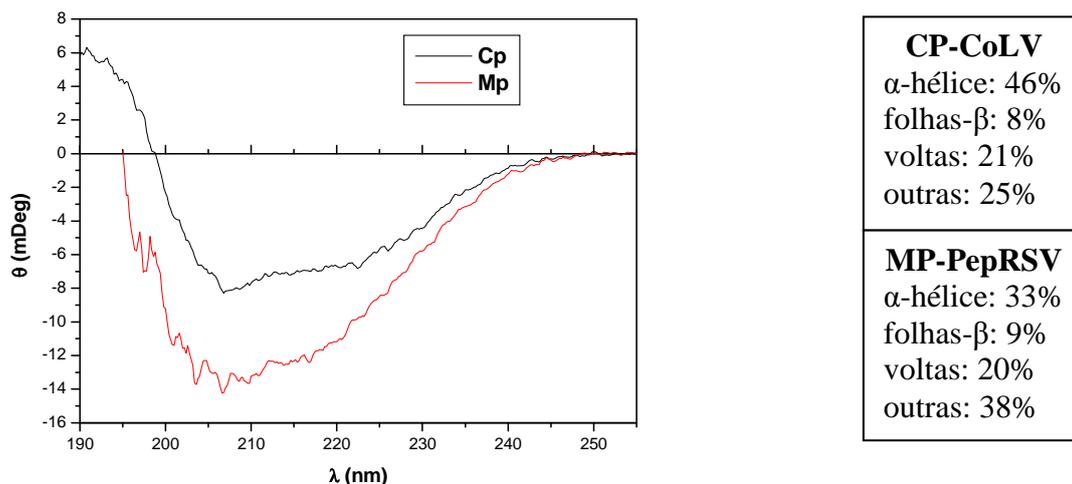
Não existe um tampão universal para o reenovelamento, é necessário testar diferentes condições e limitar o que é mais eficiente para cada proteína (YASUDA, et al., 1998). A melhor forma de reenovelamento, diálise ou diluição, não pode ser deduzida com base em propriedades moleculares das proteínas. O procedimento adequado deve ser determinado caso a caso (RUDOLPH e LILIE, 1996).

As amostras purificadas foram submetidas a vários experimentos de reenovelamento, tanto por diluição quanto por diálise. A princípio, os resultados não foram satisfatórios, havendo intensa sedimentação das proteínas. Após numerosas tentativas, encontramos um protocolo eficiente para o reenovelamento de proteínas através do método de diálise, no qual, remove-se o agente desnaturante (uréia 8 M) pela permuta em tampão fosfato/acetato pH 5,0.

Após o processo de diálise, as amostras foram concentradas até 1,0 mg/mL para serem analisadas por dicroísmo circular (CD). Obtivemos resultados apenas para a CP-CoLV e MP-PepRSV (Figura 33). A amostra de HEL do PepRSV não apresentou condições para ser analisada, provavelmente devido a baixa estabilidade da proteína.

Desde 1960 a espectroscopia por dicroísmo circular tem sido utilizada como metodologia altamente eficiente para a determinação dos diferentes tipos de estruturas secundárias em proteínas e polipeptídeos. Quando tais moléculas interagem com a luz ultravioleta provocam uma alteração nessa luz incidente e seu espectro reflete a porcentagem de cada tipo de estrutura (WHITMORE e WALLACE, 2007).

Os resultados obtidos por CD mostram a presença de 46% de alfa-hélices; 8% de folhas-beta; 21% de voltas e 25% de outras estruturas na amostra de CP-CoLV e a presença de 33% de alfa-hélices; 9% de folhas-beta; 20% de voltas e 38% de outras estruturas na amostra de MP-PepRSV. A constatação da presença de estruturas secundárias nas amostras confirma o sucesso do processo de reenovelamento dessas proteínas.



**Figura 33.** Análise por dicroísmo circular das proteínas CP-CoLV (linha preta do gráfico) e MP do PepRSV (linha vermelha do gráfico)

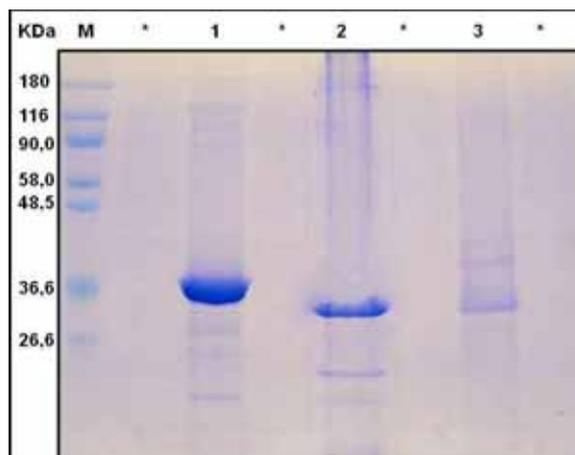
Dados estruturais sobre a proteína do capsídeo dos carlavírus não são numerosos, entretanto um estudo recente com a proteína do *Lily symptomless virus* (LSV) mostra que a CP-CoLV seria similar na composição de estrutura secundária (38% α-Helice, 10% β-folha, 4% β-volta e 48% *Coil*), o que indicaria que a estrutura desta proteína é conservada neste grupo (WANG et al., 2010).

A proteína do movimento não apresenta conservação na sequência de aminoácidos entre as quatro superfamílias que a produzem, entretanto, o alinhamento de estrutura secundária de várias sequências, indica que há convergência quanto a estrutura tridimensional. Neste mesmo estudo, observou-se que as proteínas do movimento apresentariam comumente hélices, folhas e voltas (MELCHER, 2000). O resultado do dicroísmo é condizente com este resultado quanto a alta proporção de hélices e voltas. Entretanto, o dicroísmo circular apontou para baixa porcentagem de folhas-β, o que pode indicar uma particularidade do grupo de PepRSV.

### 5.9 Concentração das proteínas e análise por DLS (Espalhamento Dinâmico de Luz)

Para que outras análises pudessem ser realizadas, bem como os testes de cristalização, seria necessário concentrar as proteínas até, ao menos,

3mg/mL. Durante as tentativas de concentração das proteínas em filtro do tipo *centricon* verificamos a perda/degradação da MP e HEL do PepRSV. Obtivemos sucesso no processo de concentração da CP do CoLV chegando a valores de até 8 mg/mL. O gel abaixo (Figura 34) mostra o resultado do processo de concentração das proteínas.



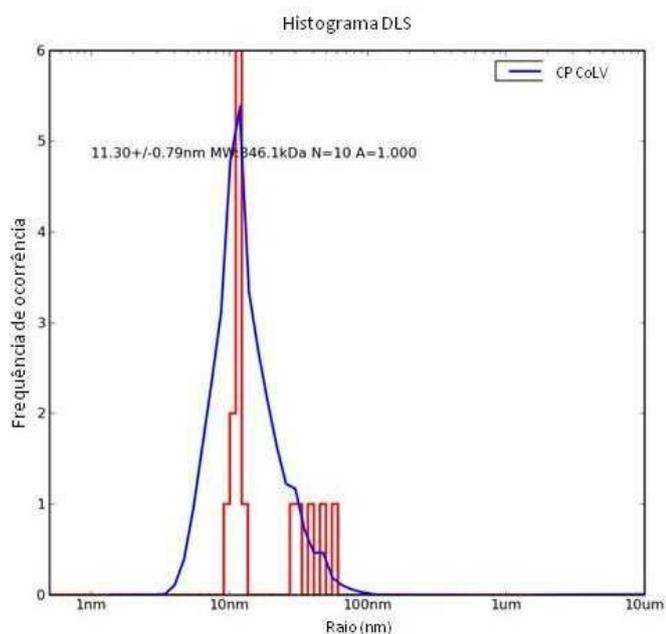
**Figura 34.** SDS-PAGE 12% mostrando o resultado do processo de concentração das proteínas reenovadas. (M) Prestained Molecular Weight Marker (Sigma); (1) CP-CoLV 3,0 mg/ml; (2) MP-PepRSV 2,0 mg/ml; (3) HEL-PepRSV 0,7 mg/ml.

Uma amostra de CP purificada (3 mg/mL) foi analisada por Espalhamento Dinâmico de Luz - DLS. Recentemente, esta técnica não-destrutiva tem obtido popularidade por explorar algumas propriedades, como tamanho e forma, de moléculas em solução (MURPHY, 1997). É uma técnica que mede a distribuição radial da solução medida e é muito usada como diagnóstico para determinar se uma amostra de proteína é adequada para cristalização pelo fornecimento de informações sobre a distribuição de tamanhos e estado de agregação da proteína em solução. Essas informações são importantes, pois estão relacionadas ao processo de cristalização, já que quanto mais homogênea (monodispersa) estiver a amostra, maior é a probabilidade de nucleação e crescimento dos cristais.

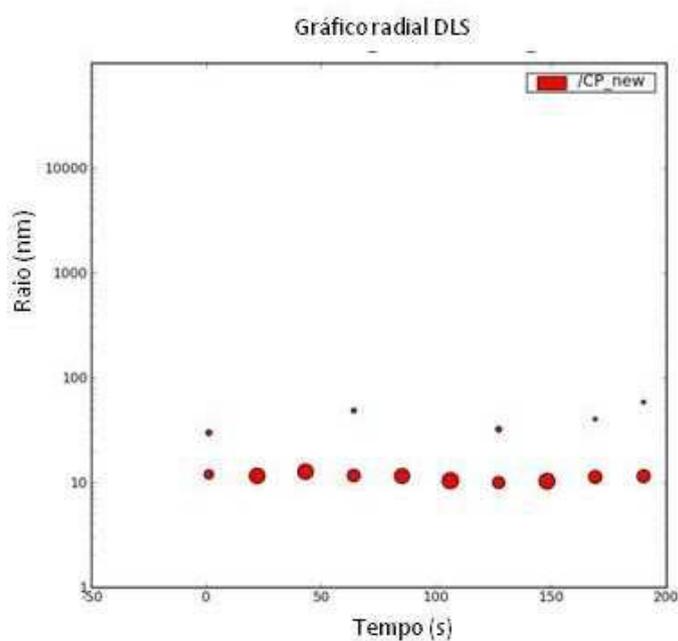
O experimento de DLS mostrou que a CP-CoLV apresenta-se estável, monodispersa e homogênea, condições que são consideradas ideais para cristalização. No gráfico 1 é possível verificar que o pico de frequência

refere-se a um raio de cerca de 10nm, esse valor relativamente baixo evidencia que não há ou há pouca agregação das proteínas.

Devido à baixa concentração e alta agregação nas amostras das outras duas proteínas, não foi possível realizar análises de DLS.



**Gráfico 1.** A abundância relativa de cada fração radial presente na amostra de CP-CoLV purificada.



**Gráfico 2.** Distribuição radial da partícula em função do tempo.

Os dados obtidos nas análises de CD e DLS mostraram que a CP do CoLV apresenta condições favoráveis ao processo de cristalização.

### 5.10 Ensaio de cristalização

Testes de estabilidade/solubilidade podem ser feitos para facilitar a determinação dos reagentes e concentrações mais adequados para determinada proteína (CAPELLE et al., 2007). Conforme sugerido por JANCARIK et al. (2004), realizamos um teste de otimização de solubilidade com 24 diferentes tampões. Foram testados 16 tipos de reagentes e variação de pH de 3,0 a 10,0. Em 3 das 24 soluções a proteína permaneceu solúvel, em 7 soluções houve precipitação fraca e nas demais soluções houve formação de forte precipitado. As soluções e a solubilidade da CP-CoLV em cada solução estão descritas na Tabela 5.

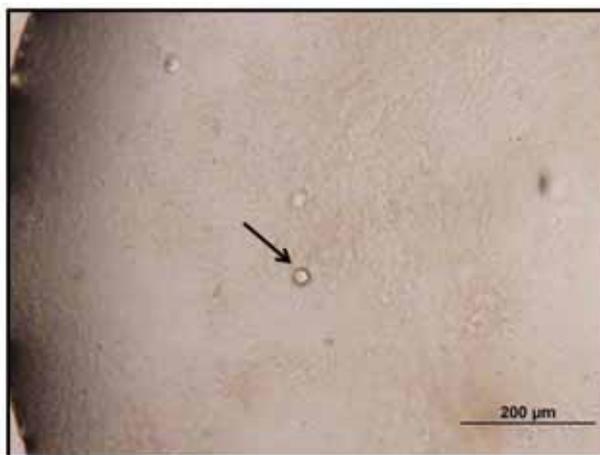
O resultado do teste de solubilidade mostrou que a CP do CoLV permanece solúvel na presença dos reagentes: glicina, acetato de sódio e MES em pH ácido variando de 3,0 a 5,8.

**Tabela 5.** Soluções utilizadas no teste de solubilidade e análise da presença de precipitado.

Condições	Presença de precipitado	Condições	Presença de Precipitado
1. Glicina pH 3,0 0,1M	<b>A</b>	13. MES pH 6,5 0,1M	<b>B</b>
2. Ácido cítrico pH 5,2 0,1M	<b>B</b>	14. Cacodilato de sódio pH 6,5 0,1M	<b>C</b>
3. Ácido cítrico pH 4,0 0,1M	<b>C</b>	15. ADA pH 6,5 0,1M	<b>C</b>
4. Acetato de sódio pH 4,0 0,1M	<b>A</b>	16. MOPS pH 7,0 0,1M	<b>C</b>
5. Acetato de sódio pH 5,0 0,1M	<b>B</b>	17. EPPS pH 8,0 0,1M	<b>C</b>
6. Citrato de sódio pH 4,7 0,1M	<b>B</b>	18. Imidazol pH 8,0 0,1M	<b>C</b>
7. Citrato de sódio pH 5,5 0,1M	<b>B</b>	19. Tris-HCl pH 8,5 0,1M	<b>C</b>
8. Fosfato de sódio pH 5,0 0,1M	<b>B</b>	20. Bicina pH 8,5 0,1M	<b>C</b>
9. Fosfato de sódio pH 6,0 0,1M	<b>C</b>	21. Bicina pH 9,0 0,1M	<b>B</b>
10. Fosfato de sódio pH 7,0 0,1M	<b>C</b>	22. CAPS pH 10,0 0,1M	<b>C</b>
11. MES pH 5,8 0,1M	<b>A</b>	23. HEPES pH 7,0 0,1M	<b>C</b>
12. MES pH 6,2 0,1M	<b>B</b>	24. HEPES pH 7,5 0,1M	<b>C</b>

**A: ausência de precipitado; B: precipitação fraca; C: precipitação forte**

Ensaio de cristalização foram realizados, utilizando a proteína CP-CoLV reenovelada por diálise. Embora não tenha obtido um cristal adequado para análises estruturais, obtivemos indícios de formação de cristal em ensaios utilizando o tampão 1 da Tabela 5 (Figura 35).



**Figura 35.** Fotomicrografia do cristal irregular, resultante do ensaio de cristalização da CP-CoLV.

### 5.11 Análise estrutural da CP-CoLV pelo PredictProtein

A predição de estrutura secundária para a CP-CoLV foi similar aos resultados de dicroísmo circular: 41,97% de hélice, 1, 64% de folha e 56, 39% de voltas e *coil*. As folhas em hélice estariam distribuídas por toda a extensão da proteína (Figura 37). Quanto a acessibilidade dos resíduos, o programa indica que a maior parte da proteína (62,95%) está exposta ao solvente em solução aquosa. Isso indica, que a proteína apresenta-se de forma pouco compacta e acessível à interação com proteínas ou moléculas na superfície.

A localização da proteína é principalmente no citosol, tanto em eucariotos quanto em procariotos. A predição de sítios de modificação pós-traducional, indica que esta proteína possui sítios de glicosilação (posição 276 e 281), fosforilação (252, 8, 82, 192, 264, 291), miristoilação (39, 79, 174) e amidação (32).

A glicosilação é uma modificação comum e é importante principalmente para interações proteína-proteína. Esta modificação não foi descrita em capsídeo de carlavírus. Em outros vírus, a glicosilação do capsídeo é importante para a transmissão e manutenção da forma funcional do capsídeo (ALTMANN e LOMONOSSOFF, 2000). A fosforilação do capsídeo também pode ser essencial ao ciclo viral. Estudos com o *Cauliflower mosaic virus*, mostra que mutações nos sítios desta modificação inibem a infecctividade do vírus (CHAPDELAINE et al., 2002). A adição do grupo miristoil e de carboidrato (ex. amidação) no capsídeo também não é descrita para os carlavirus. Entretanto, para outros vírus sabe-se que a miristoilação apresenta papel importante na entrada do vírus na célula hospedeira (MAURER-STROH e EISENHABER, 2004) e a amidação é importante para definir a antigenicidade em *Cowpea mosaic virus* (MIERMONT et al., 2002). Sabe-se que a CP em Carlavírus é responsável pela transmissão do vírus independente de proteína auxiliar (GRAY e BANERJEE, 1999). Desta forma, as modificações pós-traducionais podem ser importantes para o sucesso da infecção, participando de diferentes etapas como a interação com o vetor, a interação com o hospedeiro e a antigenicidade.

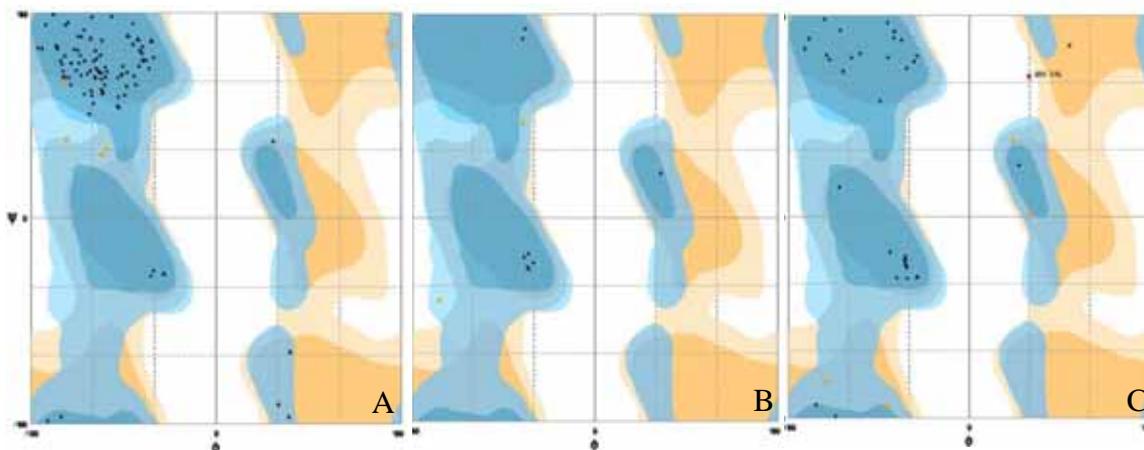
O modelo tridimensional da proteína completa foi montado utilizando as estruturas descritas na Tabela 6 como molde, identificadas pelo GenThreader.

**Tabela 6.** Estruturas proteicas utilizadas como molde para a estrutura tridimensional do CP-CoLV.

<b>Código PDB</b>	<b>Identidade</b>	<b>Descrição</b>	<b>Região na proteína-alvo</b>	<b>Nome do modelo resultante</b>
4dox	<b>32,9%</b>	Proteína do capsídeo (parcial) <i>Papaya mosaic virus</i>	<b>99-265</b>	<b>4dox_99_265</b>
2hr0	<b>11,5%</b>	Complemento C3 beta	<b>1-98/266-305</b>	<b>2hr0_1_98/ 2hr0_266_305</b>

Os modelos separados para cada região foram avaliados quanto a qualidade. A figura 36 mostra os gráficos de Ramachandran de cada um dos três modelos. Um modelo tridimensional é considerado ótimo quando >90% dos resíduos estão na área favorável e permitida do gráfico e <2% estão na

região não-favorável. O resumo da porcentagem dos resíduos em cada região é mostrado na Tabela 7.



**Figura 36.** Gráfico de Ramachandran do modelo gerado dos resíduos 99-265 (A), 1-98 (B) e 266-305(C). As áreas em azul ou laranja escuro representam regiões favoráveis, regiões permitidas em claro e as áreas não-permitidas em branco.

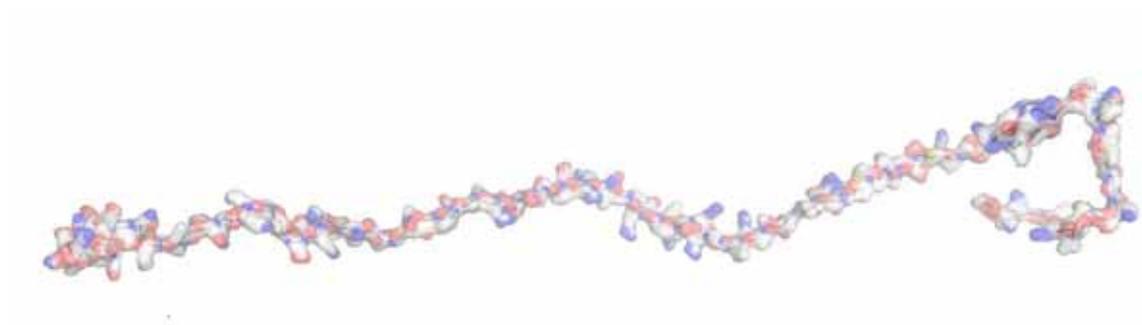
**Tabela 7.** Resumo do gráfico de Ramachandran em porcentagem dos resíduos.

Nome do modelo	Região favorável	Região permitida	Região não-permitida
4dox_99_265	93.8%	6.2%	0
2hr0_1_98	80%	20%	0
2hr0_266_305	86.8%	10.5%	2.6%

Pela avaliação de qualidade do gráfico de Ramachandran podemos concluir que os modelos foram de qualidade boa. Apenas o modelo 2hr0\_266\_305 apresentou uma quantidade de resíduos na região não permitida um pouco acima da permitida.

A avaliação do WHAT\_CHECK e do PROCHECK quanto a estereoquímica dos modelos não foram considerados satisfatórios uma vez que os valores ficaram abaixo do considerado ideal (dados não mostrados) quando comparados a proteínas com estrutura já definida. Isto pode acontecer, uma vez que modelos montados a partir de moldes com baixa identidade, podem resultar em baixa acurácia do modelo (CHAKRAVARTY et al., 2008). Entretanto, como os modelos resultantes passaram pela qualidade de ângulos da cadeia principal, estes ainda podem dar uma visão geral de como seria a

topologia da proteína (TRAMONTANO, 1998). Além disso, a dificuldade em montar um modelo rígido pode ser resultante de uma proteína com desordem natural, ou seja, com estrutura não definida e alta flexibilidade (BROWN et al., 2010). O modelo resultante, mostrando a topologia geral é mostrado na Figura 37.



**Figura 37.** Topologia geral do modelo, mostrando a cadeia principal e a superfície da CP do CoLV.

## 6. Conclusões

- Quando expressadas em *E. coli* a 37°C as três proteínas estudadas (CP do CoLV, MP do PepRSV e fragmento da HEL do PepRSV) são produzidas em forma de corpos de inclusão;
- Em expressão realizada a 18°C a CP-CoLV foi recuperada a partir da fração solúvel;
- Foi estabelecido um método de purificação desnaturante, a partir dos corpos de inclusão, no qual as proteínas atingiram ótimo grau de pureza;
- Através de diálise foi possível renovar as proteínas;
- Apenas a CP-CoLV apresentou estabilidade e concentração adequada para análises de estrutura;
- Os resultados do dicroísmo circular (CD) e do espalhamento dinâmico de luz (DLS) permitiram concluir que a proteína renovada possui estrutura secundária e estável, monodispersa e homogênea;
- Os modelos tridimensionais por modelagem molecular indicam alta flexibilidade de estrutura.



## Anexo 2 - Sequência nucleotídica da região 3'-terminal do genoma do CoLV incluindo a ORFs 5 (CP)

1381- gtgagcatggaccaaactagagcaacaagtgaaaagagcgaacaagagaaagctgctctg  
 CP⇒ M D Q T R A T S E K S E Q E K A A L  
 1441- gagcaagagaggaagaaggctcagactgaggaggaacgcaaaggaagaacacgatgta  
 E Q E R K K A Q T E E E R K G K K H D V  
 1501- gggacagctgacgggttcgagcgcttgatcgatgatgaggcatctttggaacgtagg  
 G T A D G F E A L R S D D E A S L E R R  
 1561- ttgaatgcgctgctggaatttctgcgcgatcgtaggggtgccatccgggtgactaacccg  
 L N A L R E F L R D R R G A I R V T N P  
 1621- gggttggagactgggcgaccaatgcctttgctagcggacgacatgcgccagatcccacg  
 G L E T G R P M P L L A D D M R P D P T  
 1681- aatccgtacaacagaccctcgatcgatgcacttagtaaaactcaaaccaagggtgtgtcg  
 N P Y N R P S I D A L S K L K P R A V S  
 1741- aacaacatggcaacatcgggaagacatgatgaagatttacgtcaatcttgaagggtgggc  
 N N M A T S E D M M K I Y V N L E G L G  
 1801- gtacctactgaacatgtgcagcaagtgggtgattcaagcagttgtattctgcaaagacgct  
 V P T E H V Q Q V V I Q A V V F C K D A  
 1861- agcagctcagcattcttgaccgagaggttcattcgagtggcctgggggagcgataact  
 S S S A F L D P R G S F E W P G G A I T  
 1921- gcggtatgccgtgctcgcggtgctcaaactgatgcagaaacgctgaggagggtttgcagg  
 A D A V L A V L K R D A E T L R R V C R  
 1981- ttgtacgcacccgtgacgtggaattacatggtgacgcacaacgctccgccttctgactgg  
 L Y A P V T W N Y M L T H N A P P S D W  
 2041- gcggtatgggggtccagtatgagagtcgcttcgctgcctttgactgctttgactatggt  
 A A M G F Q Y E S R F A A F D C F D Y V  
 2101- gagaatactgctgggggtgcagccgcttgaagggctaatacaggaggccgacatcacaggaa  
 E N T A G V Q P L E G L I R R P T S Q E  
 2161- aaaattgcgcaccaatacccataaggaccttgcctttgcgtagcgcgaatcgtaaccaact  
 K I A H N T H K D L A L R S A N R N Q T  
 2221- ttcgggaaacctgagtggtggaggtgactgggtggcaaaaagtggccggagctcactagggat  
 F G N L S V E V T G G K S G P E L T R D  
 2281- tacgggaaatcgaacaacaagtga

**Anexo 2.** Sequência nucleotídica da região 3'-terminal do genoma do CoLV incluindo a ORFs 5 (CP). A sequência deduzida de aminoácidos é mostrada abaixo da sequência de nucleotídeos. Asteriscos indicam o códon de terminação das proteínas. Os códons de iniciação das ORFs (atg) estão sublinhados. A tríade Histidina–Ac.Aspartico–Treonina – (HDT), possível sítio catalítico, destacada em cinza. (*Obtido de Belintani, 2008. Relatório Científico*)

## Referências Bibliográficas

- ABDELHALEEM, M. Helicases: an overview. **Methods in Molecular Biology**, v. 587, p. 1-12, 2010.
- ADAMS, A. N.; BARBARA, D. J. Host range, purification and some properties of *Hop mosaic virus*. **Annals of Applied Biology**, v. 96, p. 201-208, 1980.
- ADAMS, M. J. et al. The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. **Archives of Virology**, v. 149, p. 1045-1060, 2004.
- AGRIOS, G. N. **Plant Pathology**. 3 ed. Academic Press, San Diego, EUA, 803pp. 1988.
- ALMEIDA, A. M. R. et al. Detection and partial characterization of a carlavirus causing stem necrosis of soybeans in Brazil. **Fitopatologia Brasileira**, v.30, p.191-194, 2005.
- ALMEIDA, A. M. R.; LIMA, J. A. A. **Princípios e técnicas de diagnose aplicados a fitopatologia**. 1 ed. Sociedade Brasileira de Fitopatologia. Brasília, DF. 186pp., 2001.
- ALTMANN, F.; LOMONOSSOFF, G. P. Glycosylation of the capsid proteins of *Cowpea mosaic virus*: a reinvestigation shows the absence of sugar residues, **Journal of General Virology**, v. 81, p. 1111-1114, 2000.
- ALZHANOVA, D. V. et al. Cell-to-cell movement and assembly of a plant closterovirus: roles for the capsid proteins and Hsp70 homolog. **EMBO Journal**, v. 20, p. 6997-7007, 2001.
- AROGUNDADE, O.; BALOGUN, S. O.; ALIYU, T. H. Effects of *Cowpea mottle virus* and *Cucumber mosaic virus* on six Soybean (*Glycine max* L.) cultivars. **Virology Journal**, v. 6, p. 220-225, 2009.
- ARRONDO, J. L. R.; GONI, F. M. Structure and dynamics of membrane proteins as studied by infrared spectroscopy. **Progress in Biophysical and Molecular Biology**, v.72, p. 367-405, 1999.
- ASURMENDI, S.; BERG, R. H.; KOO, J. C.; BEACHY, R. N. Coat protein regulates formation of replication complexes during *Tobacco mosaic virus* infection. **Proceedings of the National Academy of Sciences of USA**, v. 101, p. 415-1420, 2004.
- ATREYA, P. L. et al. Mutational analysis of the coat protein N-terminal amino acids involved in potyvirus transmission by aphids. **Journal of General Virology**, v. 76, p. 265-270, 1995.

BAGNALL, R. H.; WETTER, C.; LARSON, R. H. Differential host and serological relationships of *Potato virus M*, *Potato virus S*, and *Carnation latent virus*. **Phytopathology**, v. 49, p. 435-442, 1959.

BAKER, D.; SALI, S. Protein Structure Prediction and Structural Genomics. **Science**, V. 294, P. 93-96, 2001.

BALBAS, P. Understanding the art of producing protein and non-protein molecules in *Escherichia coli*. **Molecular Biotechnology**, v. 19, p. 251-267, 2001.

BARBIERI, M. R. et al. Expressão em *Escherichia coli* da proteína capsidial do *Watermelon mosaic virus* e produção de anti-soro. **Fitopatologia Brasileira**, v. 29, p. 215-219, 2004.

BATAS, B.; CHAUDHURI, J. B. Protein folding at high concentration using size exclusion chromatography. **Biotechnology and Bioengineering**. v. 50, p. 16-23, 1996.

BELINTANI, P. **Vírus latente da couve: caracterização molecular e localização subcelular**. Tese (Doutorado em Genética). IBILCE/UNESP-SP. 2001.

BELINTANI, P.; GASPAR, J. O.; TARGON, M. L. P. N.; MACHADO, M. A. Molecular evidences confirm *Cole latent virus* as a Carlavirus. **Journal of Phytopathology**. v. 150, p. 330-333, 2002.

BERGH, S. T. et al. The nucleotide sequence of tobacco rattle virus RNA-2 (CAM strain). **Nucleic Acids Research**, v. 13, p. 8507-8518, 1985.

BERGH, S. T.; SIEGEL, A. Intraviral homology and subgenomic RNAs of *Pepper ringspot virus*. **Virology**, v. 169, p. 339-343, 1989.

BOL, J. F. Replication of alfamo- and ilarviruses: role of the coat protein. **Annual Review of Phytopathology**, v.43, p. 39-62, 2005.

BRANDES, J.; BERCKS, R. Gross morphology and serology as a basis for classification of elongated plant viruses. **Advances in Virus Research**, v. 11, p. 1-24, 1965.

BROWN, C. J.; JOHNSON, A. K.; DAUGHDRILL, G. W. Comparing models of evolution for ordered and disordered proteins. **Molecular Biology Evolution**, v. 27, p. 609-621, 2010.

BRUNT, A. A.; KENTEN, R. H. *Cowpea mild mottle*, a newly recognized virus infecting cowpea (*Vigna unguiculata*) in Ghana. **Annals of Applied Biology**, v. 74, p. 67-74, 1973.

BUCHAN, D. W. A. et al. Scalable web services for the PSIPRED protein analysis workbench. **Nucleic Acids Research**, v. 41, p. 340-348, 2013.

CAMARGO, I. J. B.; KITAJIMA, E. W.; COSTA, A. S. Visualization of the *Pepper ringspot virus* in tomato pollen. **Journal of Phytopathology**, v. 64, p. 282-285, 1969.

CAPELLE, M. A. H.; GURNY, R.; ARVINTE, T. High throughput screening of protein formulation stability: Practical considerations. **European Journal of Pharmaceutics and Biopharmaceutics**, v. 65, p.131–148, 2007.

CARRINGTON, J. C. et al. Cell-to-cell and long distance transport of viruses in plants. **The Plant Cell**, v. 8, p. 1669-1681, 1996.

CARVALHO, C. M et al. The movement protein of Cowpea mosaic virus binds GTP and single-stranded nucleic acid *in vitro*. **Journal of Virology**, v. 78, p.1591-1594, 2004.

CAVILLEER, T. D. et al. Nucleotide sequence of the carlavirus associated with blueberry scorch and similar diseases. **Journal of General Virology**,v. 75, p. 711-720, 1994.

CAVILLEER, T. D. A new strain of *Potato carlavirus M*. **Plant disease**, v. 82, p. 98-102, 1998. 1998

CHAKRAVARTY, S. et al. Systematic analysis of the effect of multiple templates on the accuracy of comparative models of protein structure. **BMC Structural Biology**, v. 8, 2008.

CHAPDELAINE, Y. et al. Mutation of capsid protein phosphorylation sites Abolishes *Cauliflower mosaic virus* infectivity. **Journal of Virology**, v. 76, p. 11748-11752, 2002.

CLARCK, E.D. Refolding of recombinant proteins. **Current Opinion in Biotechnology**. v. 9, p. 157-163, 1998.

CLARCK, E.D. Protein refolding for industrial processes. **Current Opinion in Biotechnology**. v. 85,p. 43-93, 2003.

COHEN , J. Z.M. et al. Characterization of a distinct carlavirus isolated from Verbena. **Archives of Virology**, v. 148, p. 1007-1015,2003.

COLANGELI, R. et al. Three-step purification of lipopolysaccharide-free-polyhistidine-tagged recombinant antigens of *Mycobacterium tuberculosis*. **Journal of Chromatography**, v. 714, p. 223-235, 1998.

CORDIN, O.; BANROQUES, J.; TANNER, N. K.; LINDER, P. The DEAD-box protein family of RNA helicases. **Gene**, v. 367, p. 17-37, 2006.

COSTA, A. S.; KITAJIMA, E. W.; OLIVEIRA, A. R. Vírus do anel do pimentão:

um integrante do grupo “rattle” do fumo. 1961. (Dados não publicados).

COSTA, A. S.; KITAJIMA, E. W.; COSTA, C. L.; NAGAI, H. A couve-comum como fonte de três vírus que afetam crucíferas. **Revisão de Olericultura**, v. 12, p. 82, 1072.

CROWLE, D. R.; PETHYBRIDGE, S. J.; WILSON, C. R. Transmission of *Hop latent* and *Hop mosaic* Carlaviruses by *Macrosiphum euphorbiae* and *Myzus persicae*. **Journal of Phytopathology**, v. 154, p. 11-12, 2006.

DATAR, R. V.; CARTWRIGHT, T.; ROSEN, C. G. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. **BioTechnology**. v. 11, p. 349-357, 1993.

DAWSON, W. O.; BUBRICK, P.; GRANTHAM, G. L. Modifications of the Tobacco mosaic virus coat protein gene affecting replication, movement and symptomatology. **Journal of Phytopathology**, v. 78, p. 783-789, 1988.

DEOM, C. M.; OLIVER, M. J.; BEACHY, R. N. The 30-kilodalton gene product of Tobacco mosaic virus potentiates virus movement. **Science**, v. 237, p. 389-394, 1987.

DILL, K. A.; SHORTLE, D. Denatured states of proteins. **Annual Review of Biochemistry**. v. 60, p. 795-825, 1991.

DOGLIA, S. M. et al. Fourier transform infrared spectroscopy analysis of the conformational quality of recombinant protein within inclusion bodies. **Biotechnology Journal**, v. 3, p. 193-201, 2008.

DOLJA, V. V. et al. Distinct functions of capsid protein in assembly and movement of *Tobacco etch potyvirus* in plant. **The European Molecular Biology Organization Journal**, v. 13, p. 1482-1491, 1994.

DONG, A.; HUANG, P.; CAUGHEY, W. S. Protein secondary structures in water from 2<sup>nd</sup>-derivative amide-I infrared spectra. **Biochemistry**, v. 72, p. 3303-3308, 1990.

EASTWELL, K. C.; DRUFFEL, K. L. Complete genome organization of American *Hop latent virus* and its relationship to carlaviruses. **Archives of Virology**, v. 157, p. 1403-1406, 2012.

ESWAR, N. et al. Comparative protein structure modeling with MODELLER. **Current Protocols in Bioinformatics**, Supplement, v. 15, 5.6.1-5.6.30, 2006.

FAHEY, E. M.; CHAUDHURI, J. B.; BINDING, P. Refolding and purification of urokinase plasminogen activator fragment by chromatography. **Journal of Chromatography**. v. 37, p. 225-235, 2000.

FAHNERT, B.; LILE, H.; NEUBAUER, P. Inclusion bodies: formation and utilization. **Advances in Biochemical Engineering/Biotechnology**. v. 89, p. 93-142, 2004.

FERNÁNDEZ, A.; LAIN, S.; GARCIA, J. A. RNA helicase activity of the Plum pox potyvirus CI protein expressed in *Escherichia coli*. Mapping of an RNA binding domain. **Nucleic Acid Research**, v. 23, p. 1327-1332, 1995.

FERRER, M.; CHERNIKOVA, T. N.; TIMMIS, K. N.; GOLYSHIN, P. N. Expression of a temperature sensitive esterase in a novel chaperone based *Escherichia coli* strain. **Applied environmental microbiology**. V. 70, p. 4499-4504, 2004.

FIELDS, B. N. B. et al. **Fundamental Virology**, 3 ed. Ed. Raven Publishers, Philadelphia, USA, 1996.

FISCHER, B.; SUMMER, I.; GOODENOUGH, P. Isolation, renaturation and formation of disulfide bonds of eukaryotic proteins expressed in *E. coli* as inclusion bodies. **Biotechnology and Bioengineering**. v. 41, p. 3-13, 1993.

FOSTER, G. D. Molecular variation between ordinary and Andean strains of potato virus S. **Research Virology**, v. 142, p. 413-416, 1991.

FOSTER, G. D.; MILLS, P. R. Evidence for the role of subgenomic RNAs in the production of *Potato virus S* coat protein during *in vitro* translation. **Journal of General Virology**, v. 71, p.1247-1249, 1990.

FOSTER, G. D.; MILLS, P. R. Analysis of the coat proteins of the ordinary and Andean strains of potato virus S and antigenic comparisons of carlaviruses. **Acta Virology**, v. 35, p. 184-190, 1991.

FRAENKEL-CONRAT, H.; WILLIAMS, R. C. Reconstitution os active *Tobacco mosaic virus* from its inactive protein and nucleic acids components. **Proceedings of the National Academic of Sciences of USA**, v. 15, p. 690-698, 1955.

FRITSCH, C.; MAYO, M. A.; HIRTH, L. Further studies on the translation products of *Tobacco rattle virus* RNA *in vitro*. **Virology**, v. 77, p. 722, 1977.

FUJI, S. et al. Complete nucleotide sequence of the genomic RNA of *Aconitum latent virus* (genus *Carlavirus*) isolated from *Delphinium* sp. **Archives of Virology**, v. 147, p. 865-870, 2002.

GASPAR, J. O.; VEGA, J.; CAMARGO, I. J. B.; COSTA, A. S. An structural study of particle distribution during microsporogenesis in lian tomato plants infected with the brazilian *Tobacco Rattle Virus*. **Canadian Journal of Botany: Revue Canadienne de Botanique**, v. 62, p. 372-378, 1984.

GASPAR, J. O.; COSTA, A. S. Vírus do mosaico angular do feijoeiro:

purificação e ultraestrutura dos tecidos infectados. **Fitopatologia Brasileira**, v. 18, p. 534-540, 1993.

GILLESPIE, T. et al. Functional analysis of a DNA-shuffled movement protein reveals that microtubules are dispensable for the cell-to-cell movement of Tobacco mosaic virus. **The Plant Cell**, v. 14, p. 1207-1222, 2002.

GORBALENYA, A. E.; KOONIN, E. V. Helicases: amino acid sequence comparisons and structure-function relationships. **Current Opinion in Structure Biology**, v. 3, p. 419-429, 1993.

GRAMSTAT, A.; PRÜFER, D.; ROHDE, W. The nucleotide acid-binding zinc finger protein or *Potato virus M* is translated by internal as well as by ribosomal frameshifting involving a shifty stop codon and a novel mechanism of P-site slippage. **Nucleic Acids Research**, v. 22, p. 3911-3917, 1994.

GU, Z.; SU, Z.; JANSON, J. C. Urea gradient size exclusion chromatography enhance the yield of lysozyme refolding. **Journal of Chromatography A**, v. 918, p. 311-318, 2001.

HAGEDORN, D. J.; HANSON, E. W. A comparative study of the viruses causing wisconsin pea stunt and *Red clover vein mosaic*. **Phytopathology**, v. 41, p. 813-819, 1951.

HARRISON, B. D. et al. Sixteen groups of plant viruses. **Virology**, v. 45, p. 356-363, 1971.

HARRISON, B. D. *Pea early-browning virus*. **Descriptions of Plant Viruses**, n. 120, 1973.

HARRISON, B. D.; ROBERTS, I. M. Association of *Tobacco rattle virus* with mitochondria. **Journal of General Virology**, v. 3, p.121-124, 1968.

HARRISON, B. D.; ROBINSON, D. J. Tobamoviruses. **The Plant Viruses**, v. 2, p. 339-369, 1986.

HEINLEIN, M. et al. Interaction of tobamovirus movement proteins with the plant cytoskeleton. **Science**, v. 270, p. 1983-1985, 1995.

HERNÁNDEZ, C. et al. Transmission of tobacco rattle virus isolate PpK20 by its nematode vector requires one of the two non-structural genes in the viral RNA 2. **Journal of General Virology**, v. 78, p. 465-467, 1997.

HOOFT, R. W. W.; VRIEND, G.; SANDER, C.; ABOLA, E. E. Errors in protein structures. **Nature**, v. 381, p. 272-272, 1996.

HOLLINGS, M. *Chrysanthemum virus B*. **CMI/AAB Descriptions of Plant Viruses**, n 110, 1972.

HULL, R. **Matthews' Plant Virology**. 4. ed. Academic Press, San Diego, CA, USA. 1056 pp., 2002.

HULL, R. **Matthews' Plant Virology**. Elsevier Academic Press, San Diego, CA, USA. 1001 pp., 2004.

ICTV (International Committee on Taxonomy of Viruses). [www.ictvonline.org/virusTaxonomy.asp?version=2012](http://www.ictvonline.org/virusTaxonomy.asp?version=2012). Acessado em 22 de março 2013.

IVANOV K. I.; MÄKINEN K. Coat proteins, host factors and plant viral replication. **Current Opinion in Virology**, v. 2, p. 712-718, 2012.

IWAKI, M. et al. Whitefly transmission and some properties of cowpea *Mild mottle virus* on soybean in Thailand. **Plant Disease**, v. 66, p. 362-368, 1982.

JANCARIK, J. et al. Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins. **Acta Crystallographica Section**, v. 60, p. 1670-1673, 2004.

JIN, L.; PETERSON, D. L. Expression, Isolation, and characterization of the hepatitis C virus ATPase/RNA helicase. **Archives of Biochemistry and Biophysics**, v. 323, p. 47-53, 1995

JOHNSTON, I. G.; LOUIS, A. A.; DOYE, J. P. K. Modelling the self-assembly of virus capsids. **Journal of Physics: Condensed Matter**, v.22, p. 953-984, 2010.

JONES, D. T. GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. **Journal of Molecular Biology**, v. 287, p. 797-815, 1999.

KADARE, G.; HAENNI, A. Minireview: Virus-encoded RNA helicases. **Journal of Virology**, v. 71, p. 2583-2590, 1997.

KANE, J. F.; HARTLEY, D. L. Formation of recombinant protein inclusion bodies in *Escherichia coli*. **Trends in Biotechnology**. v. 6, p. 95-101, 1988.

KIEFHABER, T.; RUDOLPH, R.; KOHLER, H. H; BUCHNER, J. Protein aggregation *in vitro* and *in vivo*: a quantitative model of the kinetic competition between folding and aggregation. **Biotechnology**. v. 9, p. 825-829, 1991.

KING, A. M. Q.; LEFKOWITZ, E.; ADAMS, M. J.; CARSTENS, E. B. **Virus taxonomy**: 9th report of the international committee on taxonomy of viruses. Elsevier, San Diego. 2011.

KITAJIMA, E. W. **Morfologia das partículas do vírus do anel do pimentão e ultraestrutura dos tecidos infectados**. Tese (Doutorado em Agronomia) -

ESALQ/USP, Piracicaba, 1967.

KITAJIMA, E. W.; COSTA, A. S. Association of *Pepper ringspot virus* (Brazilian tobacco rattle virus) and host cell mitochondria. **Journal of General Virology**, v. 4, p.177-181, 1969.

KITAJIMA, E. W.; OLIVEIRA, A. R.; COSTA, A. S. Morfologia das partículas do vírus do anel do pimentão. **Bragantia**, v. 28, p.1-22, 1969.

KITAJIMA, E. W.; CAMARGO, I. J. B.; COSTA, A. S. Morfologia e aspectos intracelulares do vírus latente da couve. **Bragantia**, v. 29, p. 181-190, 1970.

KITAJIMA, E. W.; LAURITIS, 1969

KOENIG, R. et al. Natural occurrence of *Helenium virus S* in *Impatiens holstii*. **Phytopathology**, v. 106, p. 133-140, 1983.

KOEHL, P.; LEVITT, M. A brighter future for protein structure prediction. **Nature**, v. 6, p. 108-111, 1999.

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

LARSEN, R. C.; WYATT, S. D. The complete nucleotide sequence and genome organization of red clover vein mosaic virus (genus Carlavirus). **Archives of Virology**, v. 154, p. 891-894, 2009.

LASKOWSKI, R. A., MACARTHUR, M. W., MOSS, D. S., THORNTHON, J. M. Pro-check: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallization**, v. 26, p.283-291, 1993.

LAWRENCE, D. M.; RAZANOV, M. M.; HILLMAN, B. I. Autocatalytic processing of the 223 kDa protein of blueberry scorch carlavirus. **Virology**, v. 207, p. 127-135, 1995.

LAZAROWITZ, S. G.; BEACHY, R. N. Viral movement proteins as probes for intracellular and intercellular trafficking in plants. **The Plant Cell**, v. 11, p. 535-548, 1999.

LEWANDOWSKI, D. J.; ADKINS, S. The tubule-forming NSm from Tomato spotted wilt virus complements cell-to-cell and long distance movement of Tobacco mosaic virus hybrids. **Virology**, v. 342, p. 26-37, 2005.

LI, M., SU, Z. Refolding of superoxide dismutase by ion exchange chromatography. **Biotechnology Letter**. v. 24, p. 919-923, 2002.

LI, M.; SU, Z.; JANSON, J. C. *In vitro* refolding by chromatographic procedures. **Protein Expression and Purification**. v. 33, p. 1-10, 2004.

LILIE, H.; SCHWARZ, E.; RUDOLPH, R. Advances in refolding of proteins produced in *E. coli*. **Current Opinion In Biotechnology**.v. 9, p. 497-501, 1998.

LOBLEY, A.; SADOWSKI, M. I.; JONES, D. T. pGenTHREADER and pDomTHREADER: New methods for improved protein fold recognition and superfamily discrimination. **Bioinformatics**, v. 25, p. 1761-1767, 2009.

LOVE, R. A. et al. The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. **Cell**, v. 87, p. 331-342, 1996.

MAACHUPALLI-REDDY, J. et al. Effect of inclusion body contaminants on the oxidative renaturation of hen egg white lysozyme. **Biotechnology Progress**, v. 13, p. 144-150, 1997.

MACFARLANE, S. A.; BROWN, D. J. F. Sequence comparison of RNA2 of nematode-transmissible and nematode-non-transmissible isolates of pea early-browning virus suggests that the gene encoding the 29 kDa protein may be involved in nematode transmission. **Journal of General Virology**, v. 76, p.1299-1304, 1995.

MACFARLANE, S. A.; WALLIS, C. V.; BROWN, D. J. F. Multiple virus genes involved in the nematode transmission of pea early browning virus. **Virology**, v. 219, p. 417-422, 1996.

MACFARLANE, S. A. Molecular biology of the tobnaviruses. **Journal of General Virology**, v. 80, p. 2799-2807, 1999.

MCGUFFIN, L. J; JONES, D. T. Improvement of the GenTHREADER method for genomic fold recognition. **Bioinformatics**, v. 19, p. 874-881, 2003.

MACKENZIE, D. J.; TREMAINE, J. H.; STACE-SMITH, R. Organization and interviral homologies of the 3'-terminal portion of potato virus S RNA. **Journal of General Virology**, v. 70, p. 1053-1063, 1989.

MATSUMOTO, R. et al. Packaging of proteases and proteoglycans in the granules of mast cells and other hematopoietic cells. A cluster of histidines on mouse mast cell protease 7 regulates its binding to heparin serglycin proteoglycans. **Journal of Biological Chemistry**,v. 33, p. 270-301, 1995.

MATTHEWS, R. E. F. **Plant Virology**. 3 ed. Academic Press Limited, San Diego, California. pp. 835, 1991.

MAURER-STROH, S.; EISENHABER, F. Myristoylation of viral and bacterial proteins. **Trends in Microbiology**, v. 12, p. 178-185, 2004.

MAYO, M. A.; FRITSCH, C.; HIRTH, L. Translation of Tobacco rattle virus RNA *in vitro* using wheat germ extract. **Virology**, v. 69, p. 408-415, 1976.

MELCHER, U. The '30K' superfamily of viral movement proteins. **Journal of General Virology**, v. 81, p. 257-266, 2000.

MELLO, S. C. M. **Caracterização do vírus latente da couve**. Dissertação de mestrado. Universidade de Brasília. 56 pp. 1980.

MELLO, S. C. M. **Caracterização do vírus latente da couve**. Brasília, 1981. pp.55. Tese (Mestrado em Ciências-Fitopatologia) – Instituto de Ciências Biológicas, Universidade de Brasília.

MIDDELBERG, A. P. J. Preparative protein refolding. **Trends in Biotechnology**, v.20, p. 437-443, 2002.

MIERMONT, A. et al. *Cowpea mosaic virus* capsid: A Promising Carrier for the Development of Carbohydrate Based Antitumor Vaccines. **Article first published online: 22 APR 2008**.

MOROZOV, S.Y. et al. *In vitro* membrane binding of the translation products of the carlavirus 7-kDa protein genes. **Virology**, v.183, p.782-785,1991.

MOROZOV, S. Y.; SOLOVYEV, A. G. Triple gene block: modular design of a multifunctional machine for plant virus movement. **Journal of General Virology**, v. 84, p. 1351-1366, 2003.

MURPHY, R. M. Static and dynamic light scattering of biological macromolecules: what can we learn? **Current Opinion in Biotechnology**, v. 8, p. 25-30, 1997.

NAGANO, H. et al. Conversion in the requirement of coat protein in cell-to-cell movement mediated by cucumber mosaic virus movement protein. **Journal of Virology**, v. 75, p. 8045-8053, 2001.

NOGUEIRA, N. L. **Determinação do sítio intracelular de síntese do RNA do vírus do anel do pimentão através de autoradiografia a nível de microscopia eletrônica**. Tese (Doutorado em Agronomia) - ESALQ/USP, Piracicaba, 1978.

NOUEIRY, A. O.; LUCAS, W. J.; GILBERTSON, R. L. Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. **Cell**, v. 76, p. 925-932, 1994.

OLIVEIRA, A. R. **Serologia aplicada ao vírus do anel do pimentão**. Tese (Doutorado em Agronomia) - ESALQ/USP, Piracicaba, 1967.

PACHECO, B.; CROMBET, L.; LOPPNAU, P.; COSSAR, D. A screening strategy for heterologous protein expression in *Escherichia coli* with the highest return of investment. **Protein Expression and Purification**, v. 81, p. 31-41, 2012.

PADGETT, H. S. et al. Distribution of tobamovirus movement protein in infected cells and implications for cell-to-cell spread of infection. **Plant Journal**, v. 10, p. 1079-1088, 1996.

PARKER, J. M. R.; GUO, G.; HODGES, G. S. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites?. **Biochemistry**, v. 25, p. 5425-5431, 1986.

PELHAM, H. R. B. Translation of tobacco rattle virus RNAs *in vitro*: Four proteins from three RNAs. **Virology**, v. 97, p. 256-265, 1979.

PEVZNER, P.; SHAMIR, R. **Bioinformatics for biologists**. 4<sup>a</sup> Ed. The University of Chicago Press. 329 pp., 2011.

PLOEG, A. T.; ROBINSON, D. J.; BROWN, D. J. F. RNA – 2 of tobacco rattle virus encodes the determinants of transmissibility by trichodroid vector nematodes. **Journal of General Virology**, v. 74, p.1463-1466, 1993.

POUWELS, J. et al. Studies on the origin and structure of tubules made by movement protein of Cowpea mosaic virus. **Journal of General Virology**, v. 85, p.3787-3796, 2004.

RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, v. 7, p. 95–99, 1963.

RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. **Biologia Vegetal**. 6. ed. Rio de Janeiro, RJ: Guanabara Koogan, 2001.

ROBINSON, D. J.; HARRISON, B. D. *Pepper ringspot virus*. **Descriptions of Plant Viruses**, n. 347, 1989.

ROST, B.; YACHDAV, G.; LIU, J. PredictProtein: The PredictProtein Server. **Nucleic Acids Research**, v. 32, p. 321-326, 2004.

RUDOLPH, R.; LILIE, H. *In vitro* folding of inclusion body proteins. The FASEB Journal, v.10, p. 49-56, 1996.

RUPASOV, V. V.; MOROZOV, S. Y.; KANYUKA, K. V.; ZAVRIEV, S. K. Partial nucleotide sequence of *Potato virus M* RNA shows similarities to potexvirus in gene arrangement and the encoded amino acids sequences. **Journal of General Virology**, v. 70, p. 1861-1869, 1989.

SAHDEV, S.; KHATTAR, S. K.; SAINI, K. S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. **Molecular and Cellular Biochemistry**. v. 307, p. 249-264, 2008.

SCHEIN, C. H.; NOTEBORN, M. H. M. Formation of soluble recombinant

proteins in *Escherichia coli* is favored by low growth temperature. **BioTechnology**. v. 6, p.291-294, 1988.

SCHEIN, C. H. Production of soluble recombinant proteins in bacteria. **BioTechnology**. v. 7, p. 1141-1148, 1989.

SCHLEGL, R.; IBERER, G.; MACHOLD, C.; NECINA, R.; JUNGBAUER, A. Continuous matrix assisted refolding of proteins. **Journal Chromatography A**. v. 1009, p. 119-132, 2003.

SILVA, D. M. **Estudo sobre purificação, composição e algumas propriedades do vírus do anel do pimentão**. Tese (Livre Docência) - ESALQ/USP, Piracicaba, 1965.

SINGLETON, M. R.; DILLINGHAM, D. B.; WIGLEY, D. B. Structure and mechanism of helicases and nucleic acid translocases. **Annual Review Biochemistry**, v. 76, p. 23-50, 2007.

SORENSEN, H. P.; MORTENSEN, K. K. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. **Microbial Cell Factories**. v.4, p. 1, 2005.

SHEN, W. et al. Target selection of soluble protein complexes for structural proteomics studies. **Proteome Science**, v.3, p. 3-12, 2005.

SCHOLTHOF, H. B. Plant virus transport: motions of functional equivalence. **Trends in Plant Science**, v.10, p. 376-382, 2005.

SCHULTZ, E. S.; FOLSOM, D. Transmission, variation and control of certain degeneration diseases of Irish potatoes. **Journal of Agricultural Research**, v. 25, p. 43-117, 1923.

SREERAMA, N.; WOODY, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. **Analytical Biochemistry**, v. 15, p. 252-260, 2000.

STORMS, M. M. H. et al. The nonstructural NSm protein of Tomato spotted wilt virus induces tubular structures in plant and insect cells. **Virology**, v. 214, p. 485-493, 1995.

STUBBS, G. Tobacco mosaic virus particle structure and the initiation of disassembly. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 354, p. 551-557, 1999.

SWANSON, M.; BARKER, H.; MACFARLENE, S. A. Rapid vascular movement of tobnaviruses does not require coat protein: evidence from mutated and wild-type viruses. **Annals of Applied Biology**, v. 141, p. 259-266, 2002.

TRAMONTANO, A. Homology modeling with low sequence identity. **Methods**, v. 3, p. 293-300, 1998.

TSUMOTO, K. et al. Role of arginine in protein refolding, solubilization, and purification. **Biotechnology Progress**, v. 20, p. 1301-1308, 2004.

TZFIRA, T. et al. Nucleic acid transport in plant-microbe interactions: The molecules that walk through the walls. **Annual Review of Microbiology**, v. 54, p. 187-219, 2000.

URCUQUI-INCHIMA, S.; HAENNI, A. L.; BERNARDI, F. Potyvirus proteins: a wealth of functions. **Virus Research**. v. 74, p. 157-175, 2001.

VALAX, P.; GEORGIU, G. Molecular characterization of beta-lactamase inclusion bodies produced in *Escherichia coli*. **Biotechnology Progress**, v.9, p. 539-547, 1993.

VALLEJO, L. F.; RINAS, U. Strategy for recovery of active protein through refolding of bacterial inclusion body proteins. **Microbial Cell Factories**. v. 3, p. 2-12, 2004.

VAN BELKUM, A. et al. tRNA-like properties of tobacco rattle virus RNA. **Nucleic Acids Research**, v. 15, p. 2837-2850, 1987.

VAN LENT, J. et al. Tubular structures involved in movement of Cowpea mosaic virus are also formed in infected cowpea protoplasts. **Journal of General Virology**, v. 72, p. 2615-2623, 1991.

VAN REGENMORTEL, M. H. V. et al. **Virus Taxonomy**. Seventh Report of the International Committee on Taxonomy of Viruses. San Diego: Academic Press. 2000.

VAN REGENMORTEL, M. H. V; MAHY, B. W. J. **Desk Encyclopedia of Plant and Fungal Virology**. 1ed. Academic Press, pp. 632. 2009.

VELLIOS, E. et al. Immunogold localization of tobamovirus 2b nematode transmission helper protein associated with virus particles. **Virology**, v. 300, p.118-124, 2002.

VILLAVERDE, A.; CARRIÓ, M. M. Protein aggregation in recombinant bacteria: Biological role of inclusion bodies. **Biotechnology Letters**. v. 25, p. 1385-1395, 2003.

VISSER, P. B.; BOL, J. F. Nonstructural proteins of *Tabacco rattle virus* which have a role in nematode-transmission: Expression pattern and interaction with viral coat protein. **Journal of General Virology**, v. 80, p. 3273-3280, 1999.

XIANG, H. et al. Expression, purification, and functional characterization of a stable helicase domain from a tomato mosaic virus replication protein. **Protein Expression and Purification**, v. 81, p. 89-95, 2012.

XU, L. Z.; SÁNCHEZ, R.; SALI, A.; HEINTZ, N. Ligand specificity of brain lipid-binding protein. **The Journal of Biological Chemistry**, v. 40, p. 211-220, 1996.

XU, H.; D'AUBIN, J.; NIE, J. Genomic variability in *Potato virus M* and the development of RT-PCR and RFLP procedures for the detection of this virus in seed potatoes. **Virology Journal**, v. 7, p. 25-31, 2010.

WANG, R. et al. Expression, purification and characterization of the Lily symptomless virus coat protein from Lanzhou Isolate. **Virology Journal**, v. 7, p. 34-39, 2010.

WELLING, G. W.; WEIJER, W. J.; VAN DER ZEE, R.; WELLING-WESTER, S. Prediction of sequential antigenic regions in proteins. **FEBS Letters**, v. 188, p. 215-218, 1985.

WETTER, C.; MILNE, R. G. Carlaviruses. In: KURSTAK E. (Ed) **Handbook of plant virus infections comparative diagnosis**. Amsterdam: Elsevier/North Holland Biomedical Press. 1981. P. 695-730.

WHITMORE, L.; WALLACE, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and referende databases. **Biopolymers**, v.89, p. 392-400, 2007.

WOLF, S. et al. Movement protein of Tobacco mosaic virus modifies plasmodesmatal size exclusion limit. **Science**, v. 246, p. 377-379, 1989.

YASHIMOTO, M.; KUBIO, R. Oxidative refolding of denatured/reduced lisozyme utilizing the chaperone-like function of liposomes and immobilized liposome chromatography. **Biotechnology Progress**. v. 15, p. 480-487, 1999.

YASUDA, M. et al. Effect of additives on refolding of a denaturedprotein. **Biotechnology Progress**, v. 14, p. 601-606, 1998.

YI, G. et al. RNA binding by the brome mosaic virus capsid protein and the regulation of viral RNA accumulation. **Journal of Molecular Biology**, v. 391, p.314-326, 2009.

ZAVRIEV, S. K.; KANYURA, K. V.; LEVAY, K. E. The genome organization of potato virus M RNA. **Journal of General Virology**, v. 72, p.9-14, 1991.

ZERBINI Jr., F. M.; CARVALHO, M. G.; ZAMBOLIM, E. M. **Introdução a Virologia Vegetal**. Viçosa: Universidade Federal de Viçosa. 145pp. 2002.