



Universidade de Brasília
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Biologia Animal

GISELE ALVARENGA SIMPSON

**Isolamento e Cristalização de Estados Oligoméricos de Cistatinas
de Plantas**

Brasília, DF

2020

GISELE ALVARENGA SIMPSON

**Isolamento e Cristalização de Estados Oligoméricos de Cistatinas
de Plantas**

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Animal do Instituto de Ciências Biológicas da Universidade de Brasília, como parte dos requisitos para obtenção do título de Mestre em Biologia Animal.

Orientador: Prof. Dr. Napoleão Fonseca
Valadares

Brasília, DF

2020

Banca Examinadora

Prof. Dr. Napoleão Fonseca Valadares
Laboratório de Biofísica
Departamento Biologia Celular
Instituição: Universidade de Brasília
Presidente da Banca

Prof. Dr^a. Mariana de Souza Castro
Laboratório de Toxicologia
Departamento Ciências Fisiológicas
Instituição: Universidade de Brasília
Membro Interno

Prof. Dr. Marcelo de Macedo Brigido
Laboratório de Biologia Molecular
Departamento Biologia Celular
Instituição: Universidade de Brasília
Membro Externo

Prof. Dr^a. Renata Vieira Bueno
Laboratório de Biofísica
Departamento Biologia Celular
Instituição: Universidade de Brasília
Membro Suplente

In Memoriam de minha amada avó “Edna Marques da Silva Alvarenga”

Dedico este trabalho às pessoas mais importantes da minha vida: meu Marido (Pedro), minha Filha (Laura) e minha Família (Avô e Tios). Pessoas que fazem com que meu esforço de tentar ser cada dia melhor valha a pena.

AGRADECIMENTOS

Primeiramente, ao meu singular Orientador Dr. Napoleão Fonseca Valadares, sem o qual meu mestrado poderia não ter sido concluído. Por ser um verdadeiro modelo de cientista a ser seguido. Pela atenção e apoio durante todo o processo, além do imenso suporte, através das profundas orientações passadas de maneira tão suave para o meu aprendizado, não apenas científico, durante estes anos.

Ao meu esposo Pedro Luiz Simpson por todo apoio, amor, carinho e dedicação, e por sua ajuda, tanto direta como indireta, durante este período. Igualmente, sou grata à minha filha Laura Alvarenga Simpson, por estar tão presente em minha vida.

Ao Laboratório de Biofísica Molecular por ter me acolhido.

Aos Professores do Laboratório, Dr^a Sônia Maria Freitas e Dr. João Alexandre Barbosa que, nestes anos de convivência, me passaram muitos ensinamentos, contribuindo para meu crescimento científico e intelectual. E, também, ao Prof. Dr. Aisel Valle por tirar tantas dúvidas e ser sempre essa pessoa tão gentil.

Ao Henrique, aluno de iniciação científica, pela inestimável ajuda durante o processo de realização desta pesquisa.

Aos colegas do laboratório, Wallace e Patrícia, por estarem sempre disponíveis a ajudar nos experimentos da presente pesquisa, e a todos os demais colegas pela convivência e troca de ensinamentos. A Gideane, que teve uma participação fundamental para que eu entrasse neste Laboratório. Muita gratidão Gids!

Ao servidor Chiquinho que está sempre presente em nossas vidas.

À minha primeira professora de bioquímica e eterna amiga Juliana Alves Thomé, por me ajudar profundamente nesta pesquisa, como se fosse sua, pela

alegria compartilhada nos momentos de tristeza, empurrões, puxões de orelha, enfim, por você ser quem você simplesmente é.

Às minhas amigas Elane, Márcia e Stephane (Lora), principalmente por me ouvirem mesmo quando não entendiam nada do que eu falava, pelo apoio que vocês me dão na minha vida sempre. Lora muita gratidão a você, que compartilhou literalmente esse mestrado comigo, alegrias, depressões, surtos e mais surtos, risadas e mais risadas, pouco espaço pra escrever o tanto que você me faz bem.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro ao me conceder essa bolsa de estudo.

À FAPDF pelo apoio financeiro.

Sem o apoio de cada um de vocês esta pesquisa certamente não ocorreria.

APOIO FINANCEIRO

A dissertação de mestrado intitulada “Isolamento e Cristalização de Estados Oligoméricos de Cistatinas de Plantas” desenvolvida por Gisele Alvarenga Simpson sob a orientação do Prof. Dr. Napoleão Fonseca Valadares teve o apoio financeiro do CNPq (pela concessão de bolsa de estudo – mestrado), da FAPDF (processo nº. 193.001.668/2017), da CAPES (por meio do apoio financeiro ao programa de Pós Graduação em Biologia Animal) e da FUB–UnB.

“Distrust those who own more answers than questions”

(Rosa Montero, *The story of the translucent king*, 2005)

“Quand il ne peut plus lutter contre le vent et la mer pour poursuivre sa route, il y a deux allures que peut encore prendre un voilier: la cape (le foc bordé à contre et la barre dessous) le soumet à la dérive du vent et de la mer, et la fuite devant la tempête en épaulant la lame sur l’arrière avec un minimum de toile. La fuite reste souvent, loin des côtes, la seule façon de sauver le bateau et son équipage. Elle permet aussi de découvrir des rivages inconnus qui surgiront à l’horizon des calmes retrouvés. Rivages inconnus qu’ignoreront toujours ceux qui ont la chance apparente de pouvoir suivre la route des cargos et des tankers, la route sans imprévu”

(Henri Laborit, *Eloge de la fuite*, 1976)

“Navigare necesse, vivere non est necesse”

(Plutarco, *Vida de Pompeu*, 106–48 a. C.)

RESUMO

As cistatinas das plantas representam uma subfamília das cistatinas (Cys), também conhecidas como Fitocistatinas (PhyCys), sendo caracterizadas como proteínas inibidoras das enzimas cisteíno–proteases (Cys–Prot). A atividade proteica das PhyCys pode ser controlada por meio da oligomerização do tipo troca de domínios (*domain swap*), um evento de dimerização em que as proteínas trocam elementos de estrutura secundária, formando um dímero sem atividade inibitória. As PhyCys são inibidoras de proteases exógenas e endógenas, regulando vários processos fisiológicos como: tolerância a diferentes tipos de estresses bióticos (produzidos por fungos, bactérias e insetos) e abióticos (frio, calor e sal) tornando–as interessantes alvos de pesquisa. Os estudos das PhyCys potencializam que estas sejam ferramentas biotecnológicas aplicadas na agricultura. A presente pesquisa visou selecionar, clonar, expressar, purificar e caracterizar PhyCys. Por meio de análises, de bioinformática, em banco de dados, foram avaliadas várias PhyCys de relevância econômica, resultando na escolha de duas PhyCys, uma da *Junglas regia* (NOG1) e outra do *Coffea arabica* (CAF1). As PhyCys utilizadas nesta pesquisa foram obtidas por expressão heteróloga, utilizando as cepas de *Escherichia coli* (*E. coli*) Lemo21 (DE3) cultivadas em meio auto–indutor ZYM–5052, pois essas condições apresentaram melhor expressão solúvel de NOG1 e CAF1. As proteínas foram purificadas por cromatografia de afinidade a metal e cromatografia de exclusão molecular (SEC). As purificações de NOG1 resultaram em maior quantidade de dímeros, contudo foram observados também monômeros, e para a CAF1 foram observados principalmente monômeros. A SEC foi utilizada em ensaios de estabilidade térmica e nos ensaios de oligomerização das PhyCys. As PhyCys foram incubadas em diferentes temperaturas (T) e as duas PhyCys se mostraram estáveis em gelo. Após incubação por diferentes períodos em intervalos discretos de 20 °C a 70 °C foi possível monitorar a conversão da forma monomérica na forma dimérica, e vice–versa. Os resultados demonstram que as mudanças no estado oligomérico das PhyCys ocorreram em função da temperatura, do tempo de incubação e da concentração. As triagens para obtenção de cristais foram realizadas com o auxílio do robô Mosquito[®] em diferentes condições, depois refinadas manualmente, e os cristais obtidos foram submetidos à coleta de dados de difração de raios–X *in–house*

em um difratômetro D8 VENTURE (Bruker). Dados obtidos a 3,1 Å de resolução provenientes de cristais monocínicos de grupo espacial C 1 2 1 da NOG1 foram resolvidos por substituição molecular. Foram observados dois dímeros por troca de domínios na unidade assimétrica.

Palavras-chave: cistatinas, cristalografia, *domain-swapping*, biotecnologia.

ABSTRACT

Plant cystatins, also known as phycystatins (PhyCys) represent a subfamily of cystatins (Cys), and are proteins that inhibit the enzymes cysteine–proteases (CysProt). The protein activity of PhyCys can be controlled through domain swap oligomerization. In the dimerization process, proteins exchange elements of secondary structure, forming a dimer without inhibitory activity. PhyCys are inhibitors of exogenous and endogenous proteases, acting in tolerance to different types of biotic stresses (produced by fungi, bacteria and insects) and abiotic (cold, heat and salt) making them interesting targets for research, and as possible biotechnological tools to be employed in agriculture. The present research aimed to select, clone, express, purify and characterize PhyCys. Several PhyCys of economic relevance were evaluated, resulting in one from *Junglas regia* (NOG1) and another from *Coffea arabica* (CAF1). In order to obtain the PhyCys used in this research, the heterologous expression process was carried out in a prokaryotic system using the pET24a expression vector. In the expression tests, the strains of *Escherichia coli* (*E. coli*) Lemo21 (DE3), grown in a self–inducing medium ZYM–5052, were more efficient, both in the production of NOG1 as well as CAF1. The proteins were purified by metal affinity chromatography and molecular exclusion chromatography (SEC). The purifications of NOG1 resulted in a greater amount of dimers, however, monomers and tetramers were also observed, and for CAF1, mainly monomers were observed. SEC was used in thermal stability tests and in PhyCys oligomerization tests. The PhyCys were incubated at different temperatures (T) and the two PhyCys were stable in ice. After incubation for different periods at discrete temperature intervals it was possible to monitor the conversion from monomer to dimer, and vice–versa. The results demonstrate that the changes in the PhyCys oligomeric state occurred as a function of temperature, incubation time and concentration. The screenings for obtaining crystals were performed with the help of the Mosquito[®] robot in different conditions, and then manually refined. The obtained crystals were subjected to *in–house* X–ray diffraction on a D8 VENTURE diffractometer (Bruker). Data obtained at 3.1 Å resolution from monoclinic crystals of space group C 1 2 1 of NOG1 were resolved by molecular substitution. Two domain–swapped dimers were observed in the asymmetric unit.

Keywords: cystatins, crystallography, domain–swapping, biotechnology.

LISTA DE ILUSTRAÇÕES

FIGURA 1. Relação estrutural entre cisteíno–proteases e as cistatinas.	22
FIGURA 2. Topologia das cistatinas. No exemplo a orizacistatina–i (pdb 1eqk).....	23
FIGURA 3. Forma oligomérica das fitocistatinas.....	24
FIGURA 4. A formação de um dímero com troca de domínio	25
FIGURA 5. Mecanismo de inibição em patógenos.....	26
FIGURA 6. Sequências das fitocistatinas da noqueira (NOG1) e do café (CAF1)	35
FIGURA 7. Mapa do plasmídeo construído em vetor pET–24a(+).	36
FIGURA 8. Cristal de NOG1. <i>Loops</i> para coletar os cristais. Difratoômetro D8 VENTURE	44
FIGURA 9. Eletroforese em SDS–PAGE (15 %) das cistatinas CAF1 e NOG1. Expressão em meio auto–indutor ZYM–5052 por 10h e 45 min	46
FIGURA 10. Eletroforese em SDS–PAGE 15% do sobrenadante da lise injetado nas colunas <i>HISTRAP</i> HP	47
FIGURA 11. Eletroforese em SDS–PAGE (15%) de diferentes eluições do <i>flowthru</i> de CAF1	48
FIGURA 12. Eletroforese em SDS–PAGE (15 %) de diferentes eluições do <i>flowthru</i> de NOG1.	49
FIGURA 13. Cromatograma de purificação por afinidade a níquel, injeção de 250 ml da proteína NOG1	50
FIGURA 14. Injeção do <i>flowthru</i> coletado previamente da proteína NOG1	50
FIGURA 15. Eletroforese em SDS–PAGE (15%) das amostras de purificação de NOG1	51
FIGURA 16. Sobreposição de cromatogramas das diferentes injeções na SEC, das amostras obtidas da purificação de NOG1 nas IMAC.	52
FIGURA 17. Sobreposição de dois cromatogramas de exclusão molecular da proteína NOG1.	53
FIGURA 18. Cromatograma de purificação por afinidade a níquel, injeção de 250 ml da proteína CAF1	54
FIGURA 19. Cromatograma de purificação por afinidade a níquel, injeção do <i>flowthru</i> coletado previamente da proteína CAF1	55
FIGURA 20. Eletroforese em SDS–PAGE (15 %) das amostras de purificação de CAF1	55
FIGURA 21. Sobreposição de cromatogramas de sec oriundo da injeção das amostras obtidas na purificação por IMAC de CAF1.....	56
FIGURA 22. Eletroforese em SDS–PAGE das amostras obtidas na purificação da CAF1	57
FIGURA 23. Cromatograma resultante do processo de calibração da coluna SUPERDEX 75 (10/300)	58
FIGURA 24. Gráfico de estabilidade térmica de dímeros de NOG1, sob variações de tempo.	61

FIGURA 25. Gráfico da influência da temperatura na formação de monômeros a partir de dímeros de NOG1	62
FIGURA 26. Sobreposição de cromatogramas de exclusão molecular em uma coluna SUPERDEX 75 (10/300) de amostras contendo dímeros de CAF1	63
FIGURA 27. Sobreposição de cromatogramas de SEC ilustrando a influência da temperatura na formação de dímeros a partir de monômeros de NOG1	64
FIGURA 28. Sobreposição de cromatogramas de SEC ilustrando a influência da temperatura na formação de dímeros a partir de monômeros de CAF1	65
FIGURA 29. Sobreposição de cromatogramas de SEC ilustrando o equilíbrio de conversão de dímeros/monômeros de NOG1, sob variações de concentrações da proteína	66
FIGURA 30. Porcentagem do sinal de absorbância correspondente ao pico de dímeros e ao pico de monômeros	67
FIGURA 31. Cristais de nog1 obtidos em 0,1 M fosfato-citrato pH 4,2, 40% v/v PEG 300.	68
FIGURA 32. Imagem de cristais do complexo NOG1+PAPAÍNA.....	69
FIGURA 33. Sobreposição dos dois dímeros por troca de domínios (cadeias AB e CD) contidos na unidade assimétrica da estrutura de NOG1.....	73
FIGURA 34. Conteúdo da unidade assimétrica da estrutura NOG1	74
FIGURA 35. Interação entre os dímeros na unidade assimétrica	74
FIGURA 36. Ângulos entre as hélices em NOG1.....	75
FIGURA 37. Imagem stéreo do mapa de densidade eletrônica	76
FIGURA 38. Imagem ortoscópica ilustrando o empacotamento cristalino da estrutura da NOG1..	77

LISTA DE TERMINOLOGIA: ABREVIATURAS, SIGLAS, SÍMBOLOS E UNIDADES

[θ]	Elipticidade molar
°	Graus
°C	Graus Celsius
Å	Ångström
aa	Aminoácidos
Abs	Absorbância
A280nm	Absorção de luz na faixa de 280 nm do espectro
APS	Persulfato de amônio
C	Concentração
Cal	Calorias
Cm	Centímetro
<i>C. arabica</i>	<i>Coffea arabica</i>
Cys	Cistatinas
Cys–Prot	Cisteíno–proteases
Da	Dalton
<i>E. coli</i>	<i>Escherichia coli</i>
FAO	<i>Do inglês Food and Agriculture Organization of the United Nations</i>
G	Gramas
h	Horas
I	Intensidade de sinal
CAIM	Cromatografia de afinidade imobilizada por metais
IPTG	<i>Do inglês Isopropyl β-D-1-thiogalactopyranoside</i>
J	Joules
<i>J. regia</i>	<i>Junglas regia</i>
L	Comprimento óptico
LB	Luria Bertani
M	Molar
Mg	Miligramas
min	Minutos
mL	Mililitros
mm	Milímetros
mM	Milimolar
MM	Massa Molecular
NC–IUBMB	<i>Do ingles Nomenclature Committee of the International Union of Biochemistry and Molecular Biology</i>
ng	Nanogramas
nm	Nanômetros
OD	Densidade Óptica

PDB	<i>Do inglês Protein Data Bank</i>
PEG	Polietileno Glicol
PhyCys	Fitocistatinas
CAF1	Fitocistatina da <i>Coffea arabica</i>
NOG1	Fitocistatina da <i>Junglas regia</i>
rpm	Rotação por minuto
S	Segundos
SDS	<i>Do inglês Sodium Dodecyl Suphate</i>
SDS–PAGE	<i>Do inglês Sodium Dodecyl Suphate – Polyacrilamyde Gel Electrophoresis</i>
SEC	<i>Do inglês Size Exclusion Chromatography</i>
T	Temperatura
TEMED	N', N', N', N'–tetrametilenodiamina
Tris	Tris (hidroximetil) – aminometano
V	Volts
UV	Ultravioleta
v/v	Relação de volume/volume
Vbar	Volume específico parcial
w/v	Relação peso/volume
α	Alfa
β	Beta
ΔH	Varição de entalpia
ΔS	Varição entropia
H	Viscosidade
θ	Elipticidade
P	Densidade
Σ	Sigma
μg	Microgramas
μL	Microlitros
μm	Micrômetros
μM	Micromolar

SUMÁRIO

1 INTRODUÇÃO	18
1.2 PROTEÍNAS	18
1.3 ENZIMAS	19
1.3.1 <i>Proteases e inibidores de proteases</i>	21
1.4 TROCA DE DOMÍNIOS	23
1.5 FISILOGIA DAS FITOCISTATINAS E BIOTECNOLOGIA	25
1.6 PLANTAS ESCOLHIDAS COMO ORGANISMOS FONTES DAS FITOCISTATINAS	26
2 JUSTIFICATIVA	28
3 OBJETIVOS	29
3.1 OBJETIVO GERAL	29
3.2 OBJETIVOS ESPECÍFICOS	29
4 MATERIAIS E MÉTODOS	30
4.1 MATERIAIS	31
4.1.1 <i>Reagentes</i>	32
4.1.2 <i>Meios de cultura</i>	32
4.1.3 <i>Soluções dos antibióticos utilizados</i>	32
4.1.4 <i>Soluções para eletroforese em gel de poliacrilamida</i>	33
4.1.5 <i>Soluções para coloração/descoloração de gel de policrilamida com Coomassie Blue</i>	34
4.1.6 <i>Soluções diversas</i>	34
4.2 MÉTODOS	35
4.2.1 <i>Triagem das fitocistatinas e análises de bioinformática</i>	35
4.2.2 <i>Síntese química e subclonagem</i>	35
4.2.3 <i>Escolha da linhagem celular de Escherichia coli e transformação das células por choque térmico</i>	37
4.2.4 <i>Ensaio de expressão</i>	37
4.2.5 <i>Análise das proteínas pela técnica SDS-PAGE</i>	38
4.2.6 <i>Purificação de proteína</i>	39
4.2.7 <i>Ensaio de oligomerização</i>	41

4.2.8	<i>Ensaio de cristalização das proteínas</i>	42
4.2.9	<i>Difração por raios-X e resolução das estruturas</i>	42
5	RESULTADOS E DISCUSSÃO	45
5.1	EXPRESSÃO	45
5.2	PURIFICAÇÃO	46
5.2.1	<i>Calibração da coluna Superdex75 (10/300)</i>	58
5.3	CARACTERIZAÇÃO DO PERFIL DE OLIGOMERIZAÇÃO	60
5.4	CRISTALIZAÇÃO	68
5.5	COLETA DE DADOS DE DIFRAÇÃO DE RAIOS-X E RESOLUÇÃO DA ESTRUTURA	70
5.6	ESTRUTURA CRISTALOGRÁFICA DA NOG1	73
5.7	EMPACOTAMENTO CRISTALINO	77
5.8	FITOCISTATINAS E BIOTECNOLOGIA	78
6	CONCLUSÃO	80

1 INTRODUÇÃO

1.1 Plantas

Desde os primórdios, as plantas estão sujeitas a uma variedade de estresses (bióticos e abióticos). Uma característica das plantas é sua natureza sésstil, assim, a necessidade de desenvolver mecanismos chave para uma rápida adaptação é essencial para sua sobrevivência (Gull *et al.*, 2019).

As plantas, quando expostas a diferentes tipos de estresses, são estimuladas a produzir respostas celulares adequadas, realizando ajustes estruturais e fisiológicos, de forma rápida e precisa, por vias de sinalização, que atuam como elo recebendo o estímulo e transferindo-o para maquinaria transcricional, gerando as respostas bioquímicas e fisiológicas apropriadas para sua subsistência (Gull *et al.*, 2019). As cascatas de reações na transdução de sinais resultam na resposta biológica final (Nelson e Cox, 2011). As plantas respondem à luz, água, temperatura, outros organismos, poluentes, *etc.* Nelas acontecem reações metabólicas complexas (coordenadas e de multipassos), onde convertem energia solar em energia química, e é pelo metabolismo que também formam nutrientes e novos componentes estruturais. As proteínas e enzimas são essenciais neste processo (Hopkins, 1999).

1.2 Proteínas

Olhando a trajetória de organização ao nível celular temos uma visão complexa e elaborada de como as proteínas são mediadoras importantes, pois processos que ocorrem nos animais, plantas, fungos ou bactérias, são mediados por proteínas (Alberts *et al.*, 2013). Para realizarem estes processos, as proteínas precisam adquirir estruturas tridimensionais definidas, o que permitirá sua participação em atividades biológicas específicas, como mediar processos intra e intercelulares, ou regular o crescimento e diferenciação celular (Dobson, 2003).

As proteínas são heteropolímeros lineares que fazem parte da estruturação e da maioria dos processos bioquímicos dos organismos (Nelson e Cox, 2011). Elas são formadas de um mesmo conjunto de 20 aminoácidos que se diferenciam por suas cadeias laterais. A combinação entre estes aminoácidos resulta

em uma ampla variedade de proteínas altamente específicas, que podem possuir atividades biológicas totalmente distintas (Lodish *et al.*, 2000).

As proteínas são sintetizadas pelo processo de tradução que ocorre no citoplasma das células (Clancy e Brown, 2008). Existem quatro níveis de organização da estrutura proteica: primária, secundária, terciária e quaternária, que diferenciam a sua forma e função (Rashid *et al.*, 2015). Pode se citar como exemplo de proteínas os anticorpos, as proteínas contráteis, os fatores de transcrição, as proteínas estruturais, as proteínas de armazenamento, as proteínas de transporte e as enzimas (Nelson e Cox, 2011).

Uma das características das proteínas é a propriedade de se oligomerizar, ou seja, formar um conjunto supramolecular (Hashimoto e Panchenko, 2010). O processo de oligomerização ocorre nos organismos eucariotos e procariotos, é visto como um fenômeno importantíssimo na evolução das proteínas (Kumari e Yadav, 2019).

A formação de um oligômero ocorre quando uma proteína, no seu estado monomérico se associa à mesma (homo–proteína) ou a outra proteína (hetero–proteína), e formará associação de duas, três ou quatro proteínas. Isso leva à formação de dímeros, trímeros e tetrâmeros, respectivamente. Estima–se que esse processo ocorre cerca de 4 vezes mais na formação de homo–proteínas que de hetero–proteínas (Hashimoto e Panchenko, 2010).

Na natureza, quando ocorre uma oligomerização de proteínas elas podem ser benignas (Jaenicke e Lilie, 2000) ou causar vários distúrbios e doenças (Kumari, 2019). Por exemplo, na doença de Alzheimer e Parkinson quando acontece um defeito de oligomerização de algumas proteínas amilóides ocorre a formação de placas (Cobb e Surewicz, 2009).

1.3 Enzimas

O fisiologista alemão Wilhelm Kühne foi o primeiro a empregar o termo enzima, no ano de 1878, sendo a denominação derivada do grego *énsimo* (*én* = em e *simo* = fermento ou levedura) (Kohler, 1971).

Todas as enzimas são proteínas (exceto raras exceções), altamente especializadas, e responsáveis por quase todas as reações bioquímicas nos seres vivos (Nelson e Cox, 2011).

As enzimas facilitam e aceleram as reações bioquímicas, por serem catalisadoras naturais (Callender e Dyer, 2015), e os aumentos de velocidade nas reações químicas alcançados por elas podem chegar a vinte ordens de magnitude (Wolfenden, 2006; Kohen, 2015). Quando se compara as taxas de reações enzimáticas realizadas no ambiente celular e na água, dentro de uma célula a taxa de reação será em milissegundos e uma reação realizada em água pode demorar de dias até alguns milhões de anos (Wolfenden e Radzicka, 1995; Wolfenden e Snider, 2001; Callender e Dyer, 2015). Se não houvesse enzimas para realizar reações de catálise, algumas destas reações poderiam levar um tempo semelhante à metade da idade da terra (Wolfenden e Snider, 2001). Olhando sob este ponto de vista, as enzimas são fundamentais para a vida (Callender e Dyer, 2015).

Na catálise, as enzimas se ligam a outras moléculas, chamadas de substrato, transformando-as quimicamente, o que resulta em um composto molecular distinto, chamado de produto. O substrato se liga apenas numa região das enzimas que é denominada como sítio catalítico ou sítio ativo. Portanto, somente moléculas que possuam uma conformação espacial adequada ao sítio ativo, conseguem estabelecer uma ligação, por isso, algumas enzimas são altamente seletivas (Nelson e Cox, 2011).

As proteases são enzimas que pertencem à subclasse 4 do grupo 3 (hidrolases) (União Internacional de Bioquímica, 1992). Contudo, existe uma enorme diversidade de ação e estrutura das proteases, desta forma elas não se enquadram facilmente no sistema geral de nomenclatura enzimática. Segundo Barrett (1994) “Algumas proteases são classificadas de acordo com: (1) a reação catalisada; (2) a natureza química do local catalítico; e (3) a relação evolutiva, conforme revelada pela estrutura”. Dessa forma, ele propôs uma classificação diferente para estas proteases que não se enquadravam na nomenclatura geral. Esta alternativa a ser considerada para classificação seria: 1º as enzimas serem agrupadas por tipo catalítico, 2º por família e finalmente por reação catalisada (Barrett, 1994).

Portanto, existem diferentes formas de classificação das enzimas. As mais recentes consideram sua forma atuação, sendo elas endopeptidases ou exopeptidases (Kidrič *et al.*, 2014). Já no sistema de nomenclatura de enzimas, que é definido pelo comitê especializado de Bioquímica e Biologia Molecular (*Nomenclature Committee of the International Union of Biochemistry and Molecular Biology*), as proteases são classificadas com base em seus mecanismos catalíticos

(<http://www.chem.qmul.ac.uk/iubmb>) e na classificação do banco de dados MEROPS (<http://merops.sanger.ac.uk/>) é realizada de acordo com o resíduo de aminoácido em seu sítio reativo, sendo as proteases divididas em: serina, cisteína, aspártico, glutâmico, treonina, liases de asparagina e metaloproteases, além das peptidases mistas (Kidrič *et al.*, 2014). Houve uma atualização na classificação das enzimas em 2018 e segundo o *ExplorEnz – The Enzyme Database* a nomenclatura é baseada no tipo de reação: oxirredutases (EC 1), transferases (EC 2), hidrolases (EC 3), liases (EC 4), isomerases (EC 5), ligases (EC 6) e translocases (EC 7).

1.3.1 Proteases e inibidores de proteases

As proteases, ou enzimas proteolíticas, estão presentes em todos os organismos vivos, e sua atividade é quebrar as ligações peptídicas entre dois resíduos de aminoácidos (Kidrič *et al.*, 2014).

A proteólise, realizada por proteases, é uma ferramenta de controle de diversos processos biológicos, como a digestão, a coagulação do sangue, defesa do hospedeiro contra agentes externos, infecção, replicação viral, cicatrização de feridas e doenças progressivas, entre outras funções biológicas (Nelson e Cox, 2011; Neurath, 1989a; Neurath, 1989b).

Os inibidores de proteases foram descritos na literatura há muito tempo (Järvinen *et al.*, 1978; Hibino *et al.*, 1980, Kominami *et al.*, 1982; Järvinen e Rinne 1982; Brzin *et al.*, 1982; Taiko *et al.*, 1984) e alguns da família das cistatinas (Cys), que são inibidores de cisteíno–proteases foram subdivididos, como os da família 1, 2 e de plantas (Abe *et al.*, 1987a).

Nas plantas as principais proteases encontradas são: as cisteíno, serino, aspártico e metaloproteases (Martinez *et al.*, 2016). Em particular, a papaína, Cys–Prot do mamão (*Carica papaya*) (**Fig. 1A**) foi a primeira descrita (Martinez *et al.*, 2012), e estão correlacionadas às cistatinas de plantas (**Fig. 1B**), que são inibidoras competitivas, atuando como pseudo–substratos dessas proteases (Benchabane *et al.*, 2010). Elas se ligam nas enzimas com a introdução de dois *loops* inibitórios e parte de sua região N–terminal, inibindo a enzima de forma reversível (Stubbs *et al.*, 1990) (**Fig. 1C**).

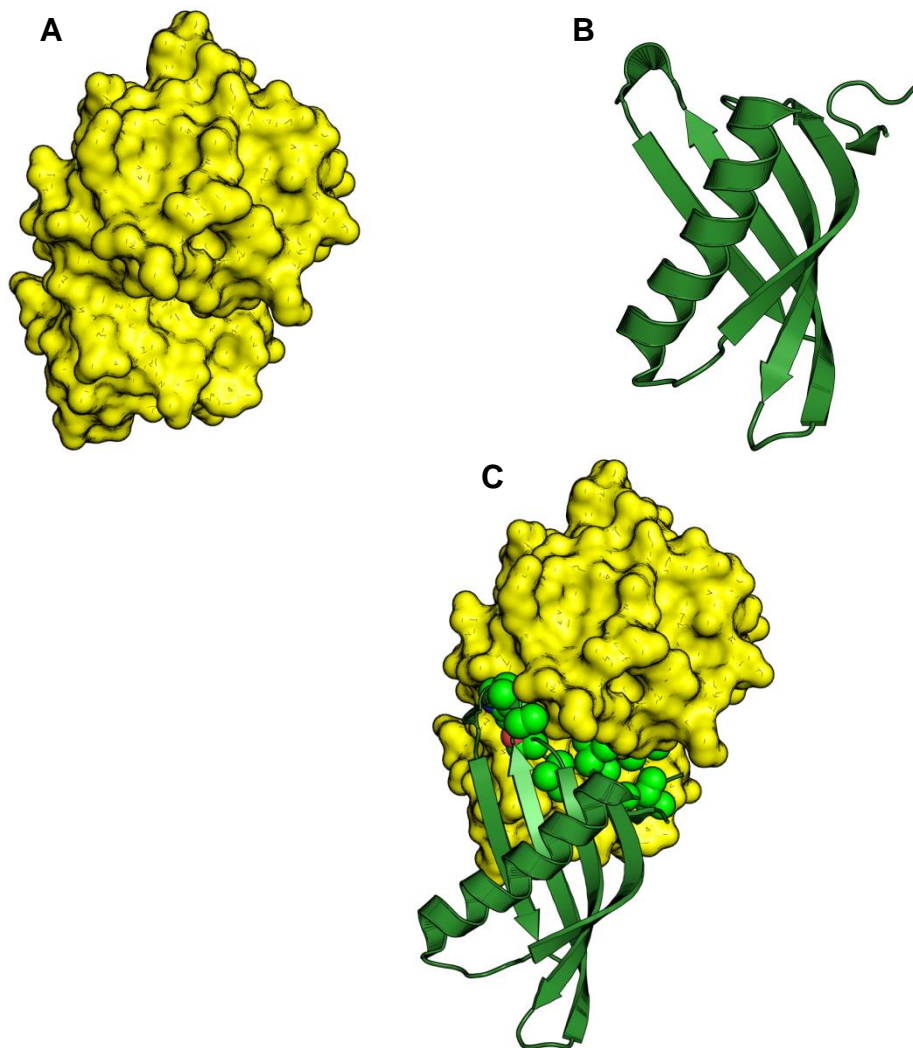


Figura 1. Relação estrutural entre cisteíno–proteases e as cistatinas. **(A)** Estrutura da papaína, uma cisteíno protease bem estudada. **(B)** Estrutura de uma fitocistatina. **(C)** Complexo formado entre a cisteíno–protease e a cistatina. Na inibição competitiva a cistatina utiliza parte de sua região N–terminal e os *loops* 1 e 2 para ocupar sítio catalítico da protease. Figuras feitas utilizando o software pymol e complexo entre a papaína e a cistatina do cará (PDB 3IMA). Apenas **(C)** está em escala.

Essas proteínas vegetais, inibidoras de Cys–Prot conhecidas como fitocistatinas, são classificadas em um agrupamento distinto de Cys (Margis *et al.*, 1998; Martinez e Dias, 2008). As PhyCys possuem relações estruturais e filogenéticas (Margis *et al.*, 1998), além de organização gênica um pouco diferente das cistatinas de animais (Kondo *et al.*, 1991). Esta subfamília, igualmente às outras famílias de cistatinas, possui a característica de inibir a atividade enzimática das

Cys–Prot (Diop *et al.*, 2004; Kordiš e Turk, 2009; Benchabane *et al.*, 2010), e foi descrita pela primeira vez por Arai e colaboradores em 1985 (Abe *et al.*, 1987b).

A estrutura terciária descrita inicialmente para a cistatina do arroz (orizacistatina) é uma α –hélice com cinco fitas β formando uma folha β antiparalela em torno desta α –hélice. Nessa estrutura a região que antecede a primeira fita β (região N–terminal) apresenta um resíduo Gly conservado, o *loop* 1 contém o motivo Gln–X–Val–X–Gly (onde X representa qualquer aminoácido), e o *loop* 2 apresenta os resíduos Pro–Trp (**Fig. 2**) (Nagata *et al.*, 2000; Turk *et al.*, 2008; Benchabane *et al.*, 2010; Martinez *et al.*, 2016).

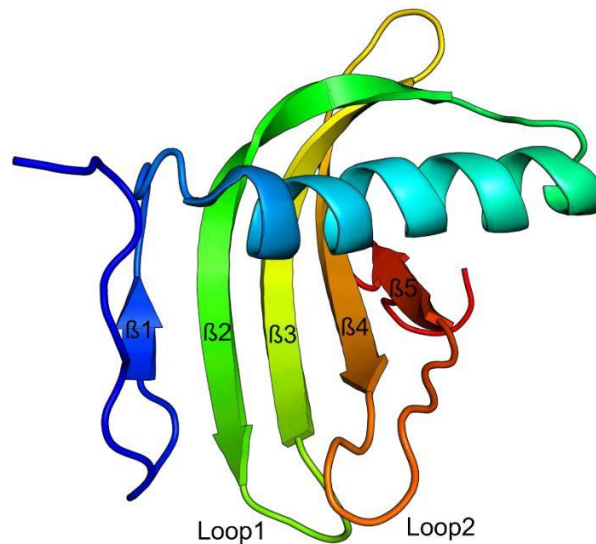


Figura 2. Topologia das cistatinas. No exemplo a orizacistatina–I (PDB 1EQK). Cinco fitas β formam a folha β antiparalela, e os *loops* estão enumerados de acordo com as convenções (NAGATA *et al.*, 2000).

1.4 Troca de Domínios

Segundo Nandwani e colaboradores (2019), o processo pelo qual proteínas monoméricas idênticas trocam elementos estruturais para gerar dímeros/oligômeros é chamado de troca de domínios (Nandwani *et al.*, 2019). Esse processo foi descrito pela primeira vez por Bennett e colaboradores em 1994. Nesse trabalho foi demonstrado que os oligômeros de uma proteína eram resultados de trocas de elementos de estrutura secundária entre monômeros, dando origem a um

dímero por troca de domínio (Bennett *et al.*, 1994). Várias proteínas se oligomerizam por troca de domínios (Bennett *et al.*, 1995).

Bennett e colaboradores (1995) definiram dois termos utilizados na troca de domínios, a “interface fechada”, que é uma região apenas observada em monômeros e que passará por uma mudança conformacional para originar a “interface aberta”, presente apenas nos dímeros por troca de domínios. No caso das cistatinas que apresentam troca de domínios, a interface aberta é formada pelos resíduos do primeiro *loop* inibitório, que conectam a fita $\beta 2$ à fita $\beta 3$ em monômeros (Júnior *et al.*, 2017) (**Fig. 3**). No caso do dímero por troca de domínios, esses resíduos do primeiro *loop* inibitório adotam uma conformação estendida, formando a interface aberta.

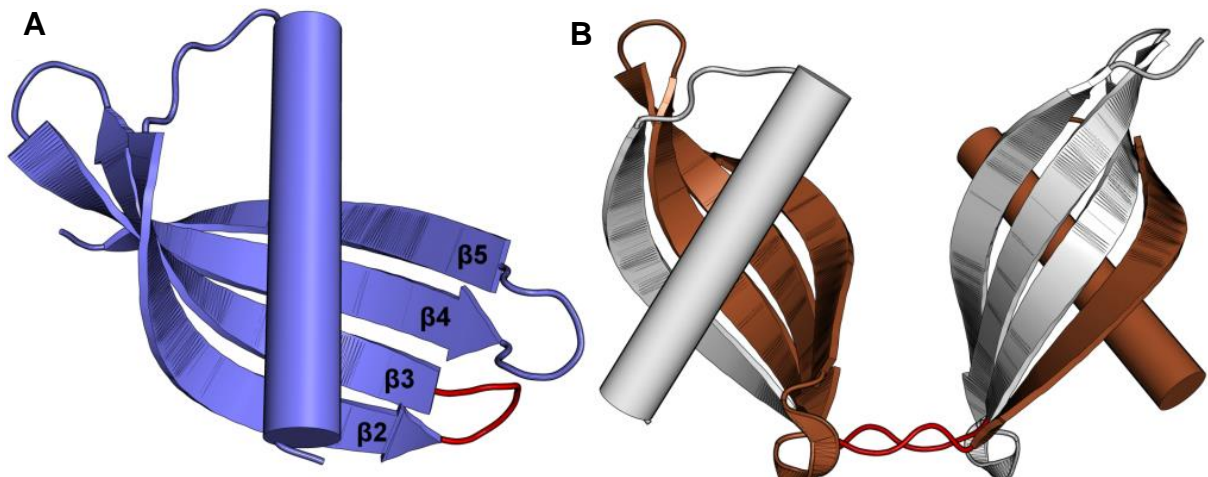


Figura 3. Forma oligomérica das fitocistatinas. Os resíduos que formam a interface aberta e a interface fechada estão coloridos em vermelho. **(A)** Fitocistatina em sua forma monomérica. **(B)** Fitocistatina na forma de dímero com troca de domínios (Júnior *et al.*, 2017).

As PhyCys possuem dois *loops* inibitórios que se ligam à protease. Conforme descrito por Valadares e colaboradores (2013) a estrutura dimérica da canecistatina–1 exibiu troca de domínios “*domain-swapped*” onde a sequência de resíduos de aminoácidos do primeiro *loop* inibitório, de um monômero, se uniu à mesma estrutura do outro monômero, fazendo com que esta conformação se tornasse inativa (**Fig. 4**) (Valadares *et al.*, 2013).

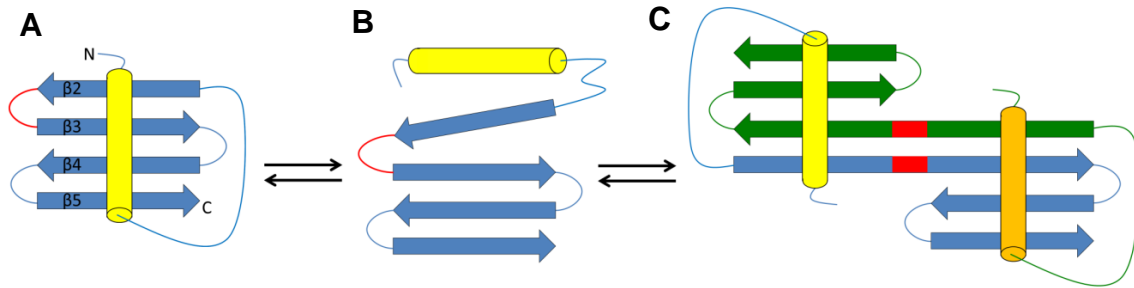


Figura 4. A formação de um dímero com troca de domínio envolve o desenovelamento parcial de dois monômeros e a subsequente montagem de um dímero utilizando elementos de estrutura secundária de duas moléculas de proteína. **(A)** Monômero. **(B)** Monômero parcialmente desenovelado. **(C)** Dímero com troca de domínios (região em vermelho) (Valadares *et al.*, 2013).

A troca de domínios em cistatinas parece estar relacionada a uma forma de regulação da atividade inibitória, pois a conversão de dois monômeros ativos em um dímero abole a atividade inibitória (Valadares *et al.*, 2013). Quando ocorre a troca de domínios formando um dímero, a topologia da proteína é conservada. Então, o dímero (estado inativo) seria uma forma de regulação fisiológica para a atividade da proteína, que pode retornar ao estado monomérico (ativo) sem a necessidade de mutações (Bennett e Eilsenberg, 2004).

1.5 Fisiologia das Fitocistatinas e Biotecnologia

As PhyCys regulam a atividade de Cys–Prot endógenas, envolvidas no desenvolvimento das plântulas e na germinação das sementes (Müntz, 1996; Toyooka *et al.*, 2000; Hwang *et al.*, 2009; Szewińska *et al.*, 2016), no desenvolvimento de nódulos nas plantas leguminosas (Naito *et al.*, 2000) e na degradação das proteínas foliares (Diop *et al.*, 2004). Além disso, regulam a atividade de Cys–Prot exógenas, como as presentes no trato digestório dos insetos herbívoros (**Fig. 5**) (Arai *et al.*, 2002; Goulet *et al.*, 2008; Martinez *et al.*, 2016), e apresentam atividade antifúngica (Yang e Yeh, 2005).

A literatura já vem descrevendo abordagens biotecnológicas do uso das PhyCys em plantas transgênicas, como por exemplo, a superexpressão ectópica delas e o impedimento dos insetos de se alimentarem destas plantas (Kiggundu *et al.*, 2010; Quain *et al.*, 2014; Martinez *et al.*, 2016), outras abordagens denotam o seu uso em substituição aos pesticidas, pela proteção que as PhyCys confere às

plantas, como contra vários tipos de pragas e estresse abiótico (Kurnet *et al.*, 2015; Hwang *et al.*, 2009). Ensaio realizados com planta transgênica (*Nicotiana tabacum*) apresentaram resistência a larvas de *Helicoverpa armigera*, a bactérias e *Oomycota* (Senthilkumar *et al.*, 2010).

No campo da biotecnologia, entender os elementos estruturais das PhyCys, bem como ocorre a regulação das Cys–Prot por mecanismos como o *domain–swapped* das PhyCys, se torna uma ferramenta valiosa.

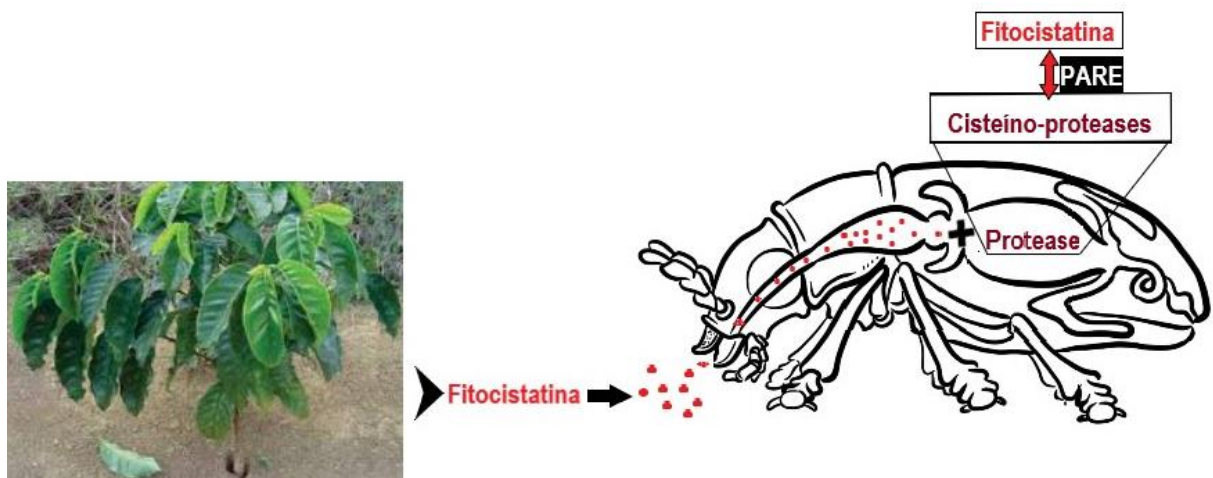


Figura 5. Mecanismo de inibição de proteases em patógenos pelas fitocistatinas presentes na árvore de café. Fitocistatinas presentes nas folhas (representadas por bolinhas vermelhas) são ingeridas pelos insetos herbívoros, no intestino médio dos insetos (mesenteron) existem as cisteíno–proteases que ao entrarem em contato com a fitocistatinas se ligam a elas, e, enquanto ligadas, as enzimas não exercem mais sua função de quebrar proteínas, desregulando, assim, a digestão do inseto.

1.6 Plantas Escolhidas como Organismos Fontes das Fitocistatinas

As PhyCys selecionadas para o presente estudo são oriundas de produtos agrícolas importantes, tanto nacionalmente, quanto internacionalmente, e o conhecimento sobre estas proteínas pode contribuir para o melhoramento genético. Para a presente pesquisa foram utilizadas PhyCys de duas espécies arbóreas, uma da *Junglas regia* (NOG1) e outra da *Coffea arabica* (CAF1).

A *Junglas regia* popularmente conhecida como noqueira–europeia, produz o fruto de onde se extrai a sua noz, muito consumida no período natalino (Backes e Irgang, 2009). Este fruto é altamente consumido em vários países por seus benefícios nutricionais. Sua madeira é nobre, e produzida no mundo inteiro

(Pollegioni *et al.*, 2014). A noqueira pertence à família Juglandaceae, nativa da Ásia central (Dandekar *et al.*, 2005) e na pré-história se espalhou até a Europa (Berry, 1912). É produzida em países como Estados Unidos e China. O Brasil é um grande importador desta noz (Martins *et al.*, 2019). Pela importância nutricional e grande consumo mundial deste produto, esta árvore possui uma grande relevância econômica e proteger esta espécie de pragas melhorando a produtividade agrícola desta cultura.

O *Coffea arabica*, uma das espécies conhecidas como café. Dos seus grãos é extraída a bebida café, consumida no mundo inteiro (CECAFE, 2019). O *Coffea* é um gênero pertencente à família Rubiaceae (Bridson e Verdcourt, 1988), originada nas florestas montanhosas do sudoeste da Etiópia, cultivada por árabes e depois por turcos que a levaram para a Europa. No século XVII, comerciantes holandeses e franceses, trouxeram a planta para América do Sul que atualmente é sua maior cultivadora (Krist, 2020). Cerca de 70% do mercado mundial do café é da *Coffea arabica* (Fazuoli, 1986).

O Brasil em 2019 foi o maior produtor, exportador e segundo maior consumidor de café mundial (CECAFE, 2019). Somente no mês de janeiro de 2020 o Brasil exportou 3,2 milhões de sacas de 60 Kg, sendo 2,7 milhões do *Coffea arabica*, este produto foi exportado com o valor médio de US\$ 138,60 por saca, obteve uma receita de US\$ 371,54 milhões (Embrapa, 2020). Considerando a receita cambial com as exportações de café brasileiro, apenas deste mês, demonstra-se que o melhoramento deste produto é importante para nosso país (CECAFE, 2020). Portanto, conhecer os mecanismos de regulação das PhyCys desta planta, que possui alta relevância econômica e conhecer possíveis mecanismos para a proteção desta espécie de pragas, garantindo uma melhor produtividade desta commodity agrícola favorece o desenvolvimento econômico do Brasil.

2 JUSTIFICATIVA

Com as mudanças climáticas e o crescimento populacional, a manutenção de produtos agrícolas para oferta de alimentos é essencial. Desta forma o melhoramento genético das plantas favorece a subsistência humana (Campos, 2019).

Fungos, insetos e outros microorganismos estão continuamente provocando danos a várias espécies de plantas. E, naturalmente, elas já possuem um mecanismo inato de proteção, que lhes garante o desenvolvimento e o sucesso reprodutivo (Carlini e Grossi de Sá, 2002). Um aparato de defesa compreende a utilização de inibidores de proteases (Dutt *et al.*, 2010; Lima *et al.*, 2015; Martinez *et al.*, 2016). Atualmente, tem-se como objetivo alcançar alta produtividade das *commodities* agrícolas. Portanto o melhoramento destas, com o controle de pragas e patógenos, além de aumentar a resistência a fatores estressores, bióticos e abióticos, visa melhorar o sistema produtivo e reduzir perdas econômicas advindas destes organismos. Desta forma, o emprego de proteínas com atividade biológica é uma maneira de minimizar os efeitos ocasionados pelo uso de insumos químicos como inseticidas e pesticidas. Isto se justifica, pois, no Brasil e no mundo, a ocorrência de pragas e doenças tem gerado grandes perdas econômicas.

O estudo das estruturas e mecanismos de oligomerização das PhyCys permitirá um melhor entendimento de sua interação com proteases, um conhecimento que pode auxiliar o desenvolvimento de plantas resistentes. Os mecanismos moleculares e outros fatores relacionados à dimerização por troca de domínios em cistatinas ainda não foram elucidados. O presente trabalho contribui com novos dados biofísicos e estruturais que auxiliam a compreensão da influência de fatores como a temperatura e a concentração em eventos de dimerização por troca de domínios.

3 OBJETIVOS

3.1 Objetivo Geral

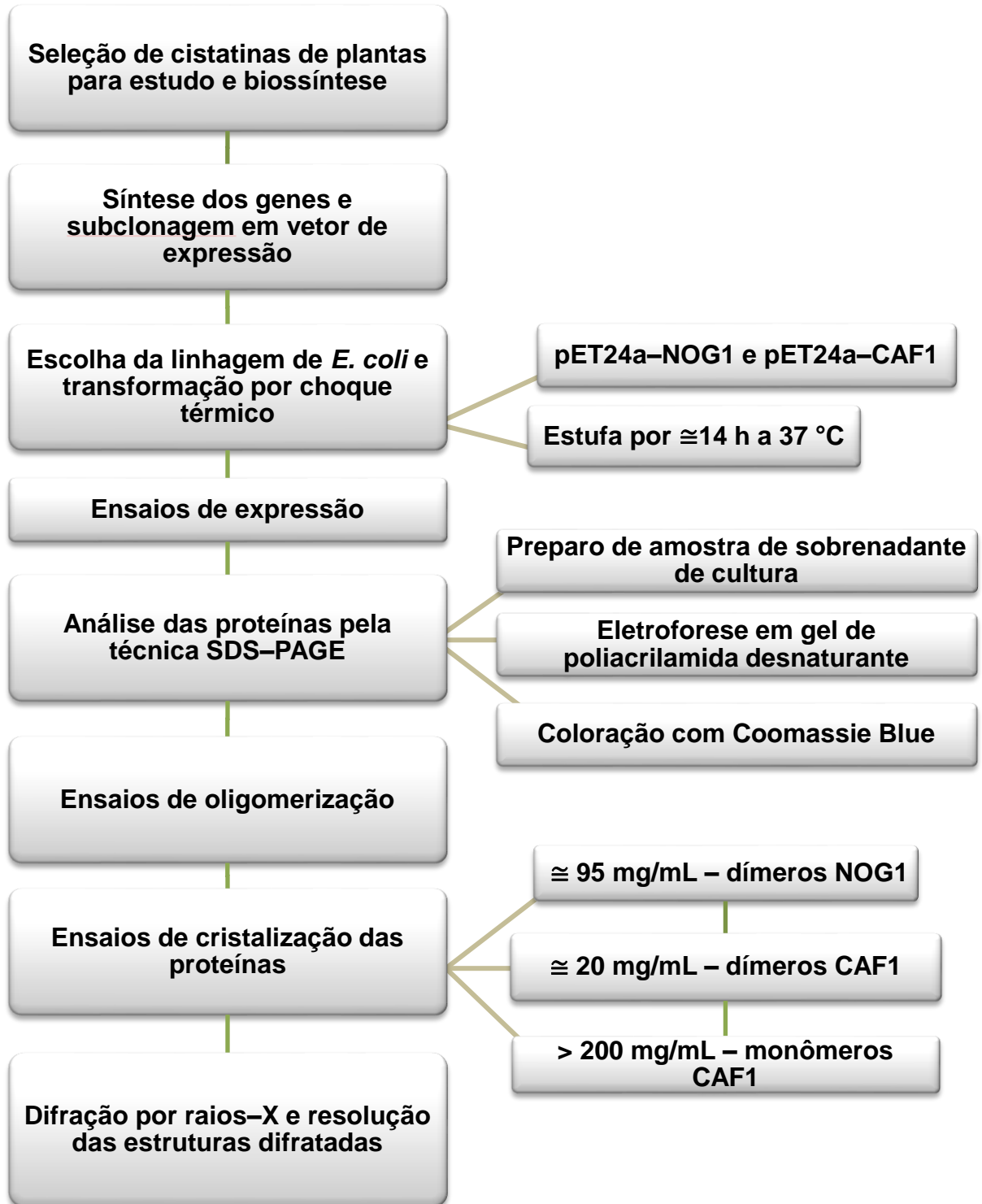
Expressar e caracterizar biofísicamente as fitocistatinas que tenham potencial biotecnológico.

3.2 Objetivos Específicos

- Seleção de cistatinas de plantas para estudo e biossíntese.
- Clonagem em vetor de expressão pET24a.
- Transformação em cepa de *E. coli* de expressão, a fim de criar os clones das proteínas escolhidas.
- Expressar as proteínas em sistema de expressão heteróloga;
- Purificar as proteínas por cromatografia de afinidade a metal;
- Purificar as proteínas por cromatografia de exclusão molecular;
- Determinar a produção de monômeros e dímeros empregando-se a técnica de cromatografia de exclusão molecular.
- Determinar a estabilidade térmica empregando-se a técnica de cromatografia de exclusão molecular;
- Determinar o comportamento dos monômeros e dímeros com ensaios de oligomerização;
- Identificar e refinar condições que propiciem a formação de cristais de proteínas;
- Obter dados de difração de raios-X dos cristais das proteínas;
- Resolução das estruturas obtidas, pelo método de substituição molecular;
- Analisar as estruturas resolvidas.

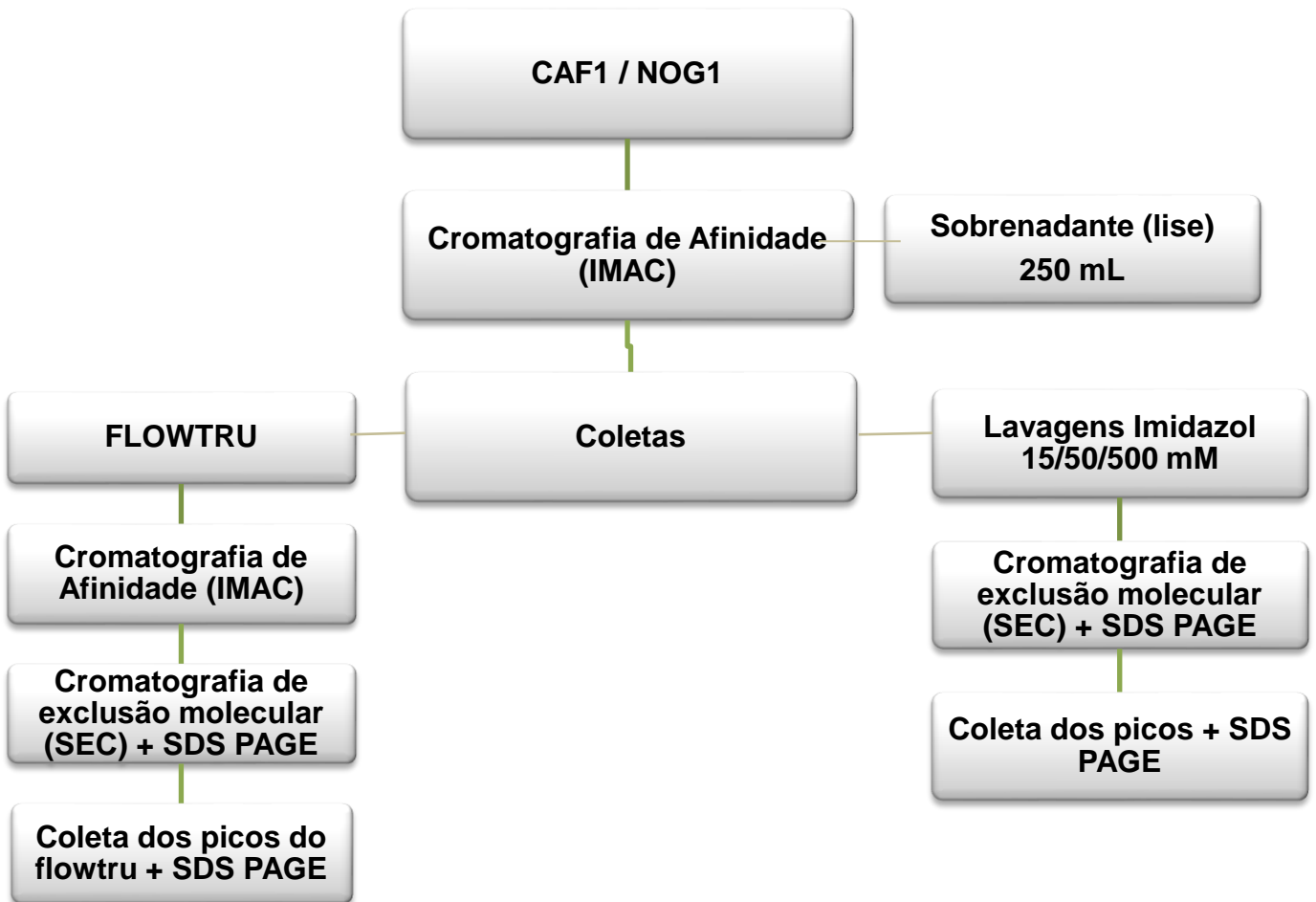
4 MATERIAIS E MÉTODOS

Organograma (fluxograma 1) da metodologia usada neste trabalho:



Fluxograma 1. Metodologia utilizada para obtenção das proteínas que foram utilizadas no presente trabalho. Obtenção por expressão heteróloga e também metodologia para ensaios realizados com estas proteínas.

Organograma (fluxograma 2) de purificações:



Fluxograma 2. Metodologia aplicada nas purificações das cistatinas obtidas por expressão heteróloga.

4.1 Materiais

Tabela 1 – Lista de equipamentos utilizados na experimentação do trabalho, presentes no laboratório de biofísica molecular da Universidade de Brasília

Material	Fabricante/País
Purificadores de Cromatografia Líquida (FLPC) Akta Primer plus e Akta Purifier	GE, EUA
Centrífuga refrigerada de alta velocidade	HITACHI, JAPÃO
Robô pipetador de placas (Mosquito®)	TTP LABTECH, EUA
Difratômetro de raios-X (D8 VENTURE)	BRUKER AXS GmbH, GER

Tabela 2 – Linhagem de bactérias para clonagem e expressão das proteínas

Linhagens	Genótipo
-----------	----------

BL21(DE3)

B F – ompT hsdSB (rB⁻ mB⁻) gal dcm (DE3)

4.1.1 Reagentes

Os reagentes de cristalização utilizados foram adquiridos da *Hampton Research*, da *Molecular Dimensions* e *Sigma*. A protease papaína foi adquirida da *Sigma*.

4.1.2 Meios de cultura

- **Meio LB (Luria–Bertani):** 1% m/v NaCl, 0,5% m/v extrato de levedura e 1% m/v peptona de caseína;
- **Meio LB sólido:** 1% m/v NaCl, 0,5% m/v extrato de levedura, 1% m/v peptona de caseína e 2% m/v ágar;
- **Meio LB KAN/CLO:** 1% m/v NaCl, 0,5% m/v extrato de levedura, 1% m/v peptona de caseína, acrescidos dos seguintes antibióticos, 100 µg/mL de Kanamicina e 34 µg/mL de Clorafenicol;
- **Meio Auto-Indutor (ZYM-5052):** 0,5% m/v extrato de levedura; 1% m/v peptona de caseína; Na₂HPO₄ 25 mM, KH₂PO₄ 25 mM, NH₄Cl 50 mM, Na₂SO₄ 5 mM, MgSO₄ 2 mM, 0,5% glicerol, 0,05% glicose, 0,2% α-lactose, FeCl₃ 10 mM.

Os meios foram preparados com água deionizada, e esterilizados por autoclavagem a 210 °C por 20 min. Depois da autoclavagem os meios de cultura foram deixados na bancada para seu resfriamento. Os antibióticos sempre foram adicionados imediatamente antes do uso, exceto no caso do meio sólido.

4.1.3 Soluções dos antibióticos utilizados

- **Solução estoque de Kanamicina (100 mg/mL):** Kanamicina foi dissolvida em água deionizada, esterilizada por filtração em membrana Millipore 0,22 µm e armazenada a -5 °C. Concentração final de uso: 100 µg/mL;

- **Solução estoque de Clorafenicol (34 mg/mL):** Clorafenicol foi dissolvido em etanol, esterilizado por filtração em membrana Millipore 0,22 µm e armazenado a -5 °C. Concentração final de uso: 34 µg/mL.

4.1.4 Soluções para eletroforese em gel de poliacrilamida

As soluções aqui descritas foram preparadas com água bidestilada ou com padrão Milli-Q e com reagentes de padrão analítico.

- **Tampão de corrida SDS:** Tris-HCl 25 mM, glicina 200 mM e SDS 0,1% (m/v);
- **Tampão para gel de proteína (separação):** Tris-HCl 1,5 M, pH 8,8. Preparação do estoque, 18,15 g de Tris dissolvido em água bidestilada 50 mL, titule com HCL para o ajuste do pH e completando com água bidestilada até ao volume de chegar 100 mL;
- **Tampão para gel de proteína (concentrador):** Tris-HCl 0,5 M, pH 6,8. Preparação do estoque, 6,05 g de Tris dissolvido em água bidestilada 50 mL, titule com HCL para o ajuste do pH e completando com água bidestilada até chegar ao volume de 100 mL;
- **Solução estoque de poliacrilamida 30% para gel de proteína:** 73 g de acrilamida, 2 g de N, N'- metileno-Bis acrilamida e 250 mL água bidestilada;
- **Solução estoque de SDS:** 10 g de SDS dissolvido em 100 mL de água bidestilada;
- **Solução de tetrametiletenodiamina (TEMED):** para um gel se usa 3 µL;
- **Solução estoque de persulfato de amônio (APS):** 0,5 g persulfato de amônio dissolvido em água bidestilada 5 mL;
- **Tampão Laemmli:** 60 mM Tris-HCl pH 6,8, SDS, 2% (m/v), β-mercaptoetanol 5% (v/v), glicerol 40% (v/v) e azul de bromofenol 0,01% (m/v).

4.1.5 Soluções para coloração/descoloração de gel de poliacrilamida com Coomassie Blue

- **Corante:** *coomassie brilliant blue R-250* 0,1% (m/v), metanol 50% (v/v) e ácido acético 10% (v/v);
- **Descorante:** 40% de metanol (v/v) e 10% de ácido acético (v/v), solução final 20% ácido acético.

4.1.6 Soluções diversas

- **Tampão padrão Tris-NaCl:** Tris-20 mM e NaCl-150 mM, pH 7,5;
- **Tampão de eluição:** **tampão padrão + 500 mM de imidazol**, pH 8,0;
- **Tampão de lise:** Tris-75 mM, NaCl-300 mM, MgCl₂-3 mM e 5% de glicerol;
- **Kit de cristalização:** Crystal Screen HR2-110 e HR2-112 (Hampton Research) I e II;
- **Kit de cristalização:** JCSG – *plus*TM (Molecular Dimensions Limited) I e II;
- **Kit de cristalização:** PACT – *premier*TM (Molecular Dimensions) I e II;
- **Kit de cristalização:** Morpheus[®] (Molecular Dimensions Limited) I e II.

4.2 Métodos

4.2.1 Triagem das fitocistatinas e análises de bioinformática

A seleção das sequências utilizadas nesse estudo priorizou cistatinas oriundas de vegetais com relevância econômica para o Brasil que possuíssem identidade sequencial relativamente alta a outras cistatinas estudadas em nosso laboratório. A premissa é que resolvendo estruturas cristalográficas de cistatinas com sequências não muito discrepantes possamos, em algum momento, definir quais resíduos são determinantes para o fenômeno de troca de domínios.

As sequências de aminoácidos de dezenas de fitocistatinas (PhyCys) foram identificadas utilizando o programa *Protein BLAST* (Altschul *et al.*, 1990). O programa *Clustal X* (Sievers *et al.*, 2011) foi utilizado para a produção de vários alinhamentos das sequências de aminoácidos de PhyCys relacionadas. Esses alinhamentos foram modificados manualmente e para cada sequência foram realizadas previsões de cristalizabilidade utilizando o servidor XtalPred–RF (<http://xtalpred.godziklab.org/XtalPred-cgi/xtal.pl>).

Por fim, foram selecionadas as sequências referentes às fitocistatinas da noqueira (NOG1) e do café (CAF1), e suas sequências estão apresentadas na **Figura 6**.



Figura 6. Alinhamento das sequências das fitocistatinas da noqueira (NOG1) e do café (CAF1). Os resíduos idênticos nas duas sequências estão destacados por uma caixa preta. A identidade destas fitocistatinas é de 66,7%.

4.2.2 Síntese química e subclonagem

A síntese dos genes das fitocistatinas da noqueira e do café e sua clonagem em pET24a(+) utilizando os sítios NdeI e XhoI foram adquiridas da empresa GenOne Biotechnologies (<http://www.genone.com.br>), de forma que recebemos em nosso laboratório os vetores pET-24a-NOG1 e pET-24a-CAF1.

No desenho do vetor também foi adicionada uma cauda de histidina (6x-His Tag) no C-terminal de cada PhyCys. Com esta 6x His Tag as PhyCys adquirem afinidade por resinas contendo NTA-Ni²⁺, que permite sua purificação pelo método cromatográfico que utiliza uma coluna de afinidade a níquel (Ni²⁺). O vetor pET-24a possui o gene lacI que codifica a proteína repressora lac, uma região de policlonagem, o promotor T7, que é específico para T7 RNA polimerase (exclusiva de eucariotos), um operador lac (bloqueador da transcrição), um gene de resistência ao antibiótico kanamicina (**Fig. 7**).

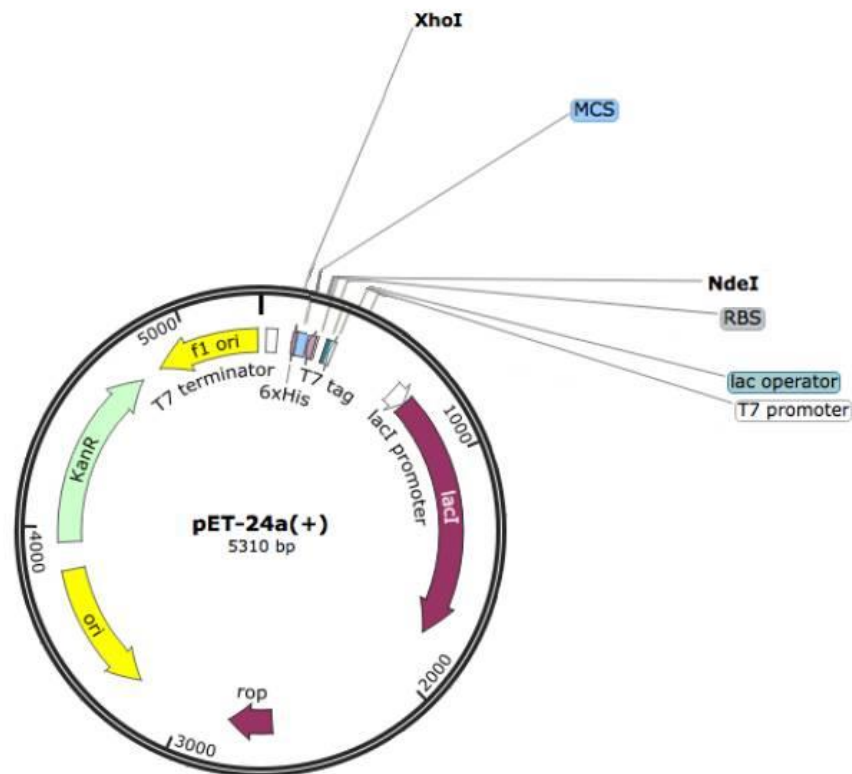


Figura 7. Mapa do plasmídeo construído em vetor pET-24a(+) com inserto de 6xHis Tag e sítios NdeI e XhoI.

4.2.3 Escolha da linhagem celular de *Escherichia coli* e transformação das células por choque térmico

As construções obtidas (pET-24a-NOG1 e pET-24a-CAF1) foram utilizadas para transformação da linhagem competente das células bacterianas, para fins da expressão heteróloga.

A transformação ocorreu em células de *E. coli* Lemo21 (DE3), e, utilizou-se $\cong 100$ ng do plasmídeo de interesse, misturando com $\cong 100$ μ L das células competentes contendo 100 mM de CaCl_2 e 15 % v/v de glicerol. A mistura foi incubada em gelo por $\cong 45$ min e sequencialmente exposta a 42 °C por 90 s. Depois deste choque térmico as células foram suplementadas com 900 μ L de meio LB e incubadas em 37 °C, sob agitação 220 rpm, por 1 h. Posteriormente, foram adicionados 500 μ L desta cultura em placas de Petri contendo meio LB sólido, e antibióticos, 100 μ g/mL de Kanamicina e 34 μ g/mL de Clorafenicol. Estas placas foram levadas para estufa por $\cong 14$ h a 37 °C.

4.2.4 Ensaio de expressão

Primeiramente é realizado um pré-inóculo contendo 5 mL de LB KLAN/CLO, neste é adicionado uma colônia isolada (retirada da placa, descrita no item anterior), são feitos três alíquotas deste procedimento e estas são incubadas sob agitação a 37 °C e 220 rpm.

4.2.4.1 Expressão em meio de auto-indução (ZYM-5052)

As expressões em grande escala foram realizadas em *Erlenmeyer* de 1 L (abaulados) contendo 500 mL de meio auto-indutor ZYM-5052, e 1 mL da cultura (descrita no item anterior).

O meio auto-indutor ZYM-5052 foi preparado conforme descrito por Studier (2005) e suplementado com os antibióticos, Kanamicina e Clorafenicol, além do FeCl_3 , que foi utilizado como fonte de íons metálicos. Os *Erlenmeyer* contendo as culturas foram incubados a 37 °C e 220 rpm.

As culturas foram transferidas para frascos específicos e estes foram centrifugados por 15 min a 4340 g, 4 °C, em uma centrífuga CR22N (*Hitachi*). O

sobrenadante foi descartado e o sedimento (*pellet*) foi congelado a $-5\text{ }^{\circ}\text{C}$, para posterior purificação das PhyCys. O mesmo procedimento SDS–PAGE foi seguido para as análises das proteínas expressas no meio ZYM–5052.

4.2.5 Análise das proteínas pela técnica SDS–PAGE

4.2.5.1 Preparo de amostra de sobrenadante de cultura

As amostras para serem analisadas em gel desnaturante foram obtidas a partir de alíquotas de 1000 μL .

Assim que obtidas as amostras, elas foram rapidamente incubadas em banho de gelo, em seguida, cada alíquota foi misturada com 30 μL de tampão *Laemmli*, e, com água deionizada (quando necessário), analisada imediatamente ou estocada a $-5\text{ }^{\circ}\text{C}$ até a análise (*a posteriori*) em SDS–PAGE (Laemmli, 1970).

4.2.5.2 Condições da eletroforese em gel de poliacrilamida desnaturante

A eletroforese das proteínas foi conduzida com gel desnaturante de poliacrilamida 15 % (m/v) imerso em tampão de corrida SDS. Amostras obtidas *a priori* foram fervidas por 5 min, para a desnaturação das proteínas, e em seguida aplicadas no gel para corrida em uma cuba de eletroforese *Mini–PROTEAN Tetra System* (BIO–RAD), ligada a uma fonte EPS 301 (GE *Healthcare*). A corrente aplicada e a tensão elétrica foram limitadas a 35 mA e 220 V, respectivamente.

Em todos os géis foi incluído um padrão de massa molecular para proteínas (*Invitrogen™*, *BenchMark™*, *Protein Ladder*).

4.2.5.3 Coloração com Coomassie Blue

Ao término da corrida eletroforética, os géis obtidos foram incubados com o corante *Coomassie Blue*, por 20 min, sob agitação e, em seguida, estes mesmos foram incubados, sob agitação, por 60 min, com a solução descolorante. Durante o tempo de descoloração dos géis, a solução descolorante foi trocada várias vezes, até possibilitar a visualização das possíveis bandas proteicas (ou não proteicas) presentes nos géis.

4.2.6 Purificação de proteína

4.2.6.1 Lise

Os pellets bacterianos congelados foram ressuspensos cada um em 45 mL tampão de lise. A lise foi feita por sonicação (ultrasonicador S250A, Branson Ultrasonics), utilizando amplitude de 20 %, em banho de gelo. Foram realizados 3 ciclos de sonicação, cada ciclo com 10 min de duração, sendo 10s de sonicação e 10 s de descanso, e amplitude 30. Depois de lisadas as amostras foram centrifugadas por 2 h a 4 °C e 17400 g o sobrenadante foi filtrado em uma membrana de poro 0,45, e suplementado com 15 mM de imidazol. Também foram alíquotadas sobrenadantes da lise.

4.2.6.2 Cromatografia de afinidade por íons metálicos imobilizados (IMAC)

A primeira etapa da purificação das PhyCys deste trabalho foi realizada no equipamento ÄKTA Primer Plus (GE *Healthcare*) com uma coluna *HisTrap*[™] HP (Ni²⁺) de 5 mL, também da GE, conectada e previamente lavada e equilibrada com tampão padrão. O sobrenadante obtido na lise foi injetado na coluna e lavagens com diferentes concentrações de imidazol foram realizadas visando remover contaminantes que interajam com a resina. Para a eluição das PhyCys foi utilizado um tampão de eluição contendo imidazol.

Em todos os procedimentos a eluição das amostras foi monitorada na faixa de absorção de luz de 280 nm. As amostras obtidas foram coletadas manualmente e armazenadas no gelo. A pureza das PhyCys de cada cromatografia de afinidade foi avaliada por SDS–PAGE.

4.2.6.3 Cromatografia de exclusão molecular (SEC)

A segunda etapa da purificação foi realizada no equipamento ÄKTA Purifier (GE *Healthcare*), onde uma coluna Superdex 75 (10/300) (GE *Healthcare*) foi conectada.

Inicialmente, esta coluna foi pré-equilibrada com o tampão padrão. Foram injetadas manualmente as amostras, obtidas no item anterior, em um *loop* de 5 mL e a eluição das proteínas ocorreu no mesmo tampão.

O fluxo foi 1 mL/min e a corrida com bombeamento isocrático de 1 volume e meio de coluna. As amostras obtidas, tanto de monômeros, quanto de dímeros, foram coletadas manualmente e armazenadas no gelo. De cada amostra coletada retirou-se uma alíquota, para confirmação por SDS-PAGE, e o mesmo procedimento foi mantido em cada SEC realizada.

4.2.6.4 Estimativa da concentração das proteínas obtidas

Para os ensaios realizados *a posteriori*, as amostras obtidas *a priori*, foram colocadas em cubetas de quartzo e as leituras foram efectuadas nos comprimentos de onda de 280 nm e 331 nm, num espectrofotómetro UV/VIS Jasco de feixe duplo, modelo V-530, utilizando a equação 1, abaixo descrita:

Equação 1:
$$A_{280} - A_{331} = X$$

Onde A_{280} é o valor obtido na leitura de absorvância em 280 e A_{331} é o valor obtido na leitura de absorvância em 331, e X é o resultado da diferença das leituras de absorvância.

O valor de X foi utilizado para calcular a concentração das proteínas segundo a lei de *Lambert Beer* (equação 2).

Equação 2:
$$C = A\lambda \cdot MW / \epsilon\lambda \cdot l$$

Onde C é a concentração em mg/mL, $A\lambda$ é o valor X (equação 1), MW é massa molecular em g/mol, $\epsilon\lambda$ é coeficiente de absorvidade molar em $M^{-1} \text{ cm}^{-1}$ e l é a distância percorrida pelo feixe luminoso através da amostra. A concentração molar determinada foi multiplicada pelo fator de diluição. Os coeficientes de extinção teóricos das proteínas foram obtidos a partir da sequência de aminoácidos das proteínas, pelo programa *Expasy* e a ferramenta *Protparam* (<http://web.expasy.org/protparam/>).

4.2.6.5 Concentração das proteínas

As PhyCys, obtidas nas purificações da SEC, foram concentradas e dialisadas, utilizando um concentrador de 4 mL com ultrafiltração de 10.000 MWCO da marca Sartorius (Sartorius™ Vivaspin™ Turbo 4 Ultrafiltration), na centrífuga com T de 4 °C e 4.000 rpm, até chegar à quantidade de mg/mL adotados nos ensaios de cristalização.

4.2.7 Ensaios de oligomerização

Os ensaios foram realizados com coluna Superdex 75 (10/300) (GE Healthcare) acoplada no equipamento ÄKTA Purifier (GE Healthcare), e equilibrada previamente com o mesmo tampão padrão. O fluxo da corrida foi 1 mL/min.

Para determinar os estados de oligomerização utilizou-se a SEC. As amostras puras de PhyCys das diferentes conformações de PhyCys, que estavam previamente armazenadas no gelo, foram expostas a diferentes tratamentos e injetadas manualmente num *loop* de 1 mL.

Os diferentes tratamentos consistiram em aquecer a amostra por um determinado tempo, assim as amostras que foram injetadas, preliminarmente passaram por várias incubações em diferentes tempos (1 min, 2 min, 5 min, 10 min, 20 min e 40 min) na temperatura (T) de 70 °C, assim se estabeleceu o protocolo de aquecimento de 20 min para investigação da mudança de estado oligomérico.

Com o tempo de 20 min de aquecimento determinado para as diferentes formas de PhyCys. Estas foram incubadas, variando a T de gelo, 20 °C, 30 °C, 40 °C, 50 °C, 60 °C, 70 °C e 80 °C, por 20 min e em seguida injetadas. Em todos estes testes, tanto a forma monomérica, quanto a forma dimérica, das diferentes PhyCys estavam no mesmo tampão padrão. E, em cada temperatura, foi utilizada uma amostra diferente da mesma conformação da proteína. Foi investigado o efeito do aquecimento, tanto para as formas diméricas, quanto para as formas monoméricas de CAF1 e de NOG1.

4.2.8 Ensaios de cristalização das proteínas

Os ensaios de cristalização foram realizados, com $\cong 95$ mg/mL, da NOG1 dímeros e $\cong 20$ mg/mL da CAF1 dímeros, $\cong 235$ mg/mL da CAF1 monômeros, além da cocristalização do complexo da papaína com a NOG1 monômeros (relação proteína/enzima 2:1). Foram utilizadas microplacas (12x8), com o conteúdo de substâncias do kit de cristalização Crystal Screen HR2-112 (I e II) e do kit de cristalização JB-JCSG (I e II).

A metodologia adotada foi a de difusão de vapor gota pendurada (*hanging drop*). As triagens foram realizadas com auxílio de um robô pipetador (*Mosquito Crystal Nanolitre Protein Crystallisation Robot, TTP Labtech, Reino Unido*) e placas de 96 poços. Apenas as NOG1 e a cocristalização do complexo da papaína com a NOG1 apresentaram possíveis formas cristalinas e seguiram para etapa de refinamento manual. Estas foram mantidas em sala de T controlada de 20 °C, sendo observadas com auxílio de uma lupa, e as condições em que cristais se formaram foram otimizadas manualmente em placas de 24 poços. Cada poço continha 500 μ L da solução mãe e cada gota continha 1 μ L do poço + 1 μ L das PhyCys. Foram realizados ensaios manuais das amostras de CAF1 (monômeros e dímeros) em placa de 24 poços.

4.2.9 Difração por raios-X e resolução das estruturas

Para a coleta de dados os cristais de NOG1 (**Fig. 8A**), e do complexo NOG1+Papaína, foram retirados das gotas com o auxílio de *loops* adequados ao tamanho dos cristais (**Fig. 8B**), e, foram submetidos à difração de raios-X, em um fluxo contínuo de nitrogênio, utilizando um difratômetro D8 VENTURE (*Bruker*) (**Fig. 8C**). O experimento de difração ocorreu no laboratório de biofísica da Universidade de Brasília, na cidade de Brasília (DF). Alguns cristais precisaram ser imersos numa solução crioprotetora (60 % solução mãe + 40 % etilenoglicol), para a crioproteção durante a difração de raios-X (McFerrin e Snell, 2002).

As imagens geradas na difração de raios-X possuem um padrão de pontos com suas respectivas intensidades. Esses pontos são chamados reflexões, e seu arranjo geométrico contém informações sobre a rede cristalina e a simetria do cristal. A informação do conteúdo da rede cristalina é obtida pelas intensidades das

reflexões. Os pontos encontrados são indexados (nomeados de acordo com o índice de Miller), e suas intensidades integradas. O procedimento de escalonamento tenta colocar todas as observações em uma escala comum.

Para processar os dados, integrar, indexar e escalonar, foram utilizados os programas: XDS (Kabsch, 2010) e o conjunto de programas CCP4 (Winn *et al.*, 2011).

O conjunto de reflexões indexadas foi dividido automaticamente em duas partes pelo programa XDSCONV, que marcou aleatoriamente 5 % das reflexões para serem utilizadas para o cálculo dos valores de R_{free} ao longo do processo de refinamento. A resolução da estrutura se deu pelo método de substituição molecular utilizando o programa PHASER do pacote CCP4 (MCCOY *et al.*, 2007). A estrutura da cistatina do lúpulo (*Humulus lupulus*) de código PDB 6VLP, que apresenta um monômero *domain swap* em sua unidade assimétrica, foi utilizada para a produção de um modelo de monômero dessa cistatina. Esse modelo foi utilizado no procedimento de substituição molecular no programa PHASER. O refinamento da estrutura foi feito utilizando os programas PHENIX 1.18.2 (Liebschner, 2019) e COOT 0.9 (Emsley, 2010). A solução do procedimento de substituição molecular foi utilizada no programa PHENIX junto com as reflexões fornecidas pelo programa XDS, e no primeiro ciclo de refinamento um procedimento de *simulated annealing* foi feito para reduzir possíveis tendências provenientes do modelo utilizado na substituição molecular.

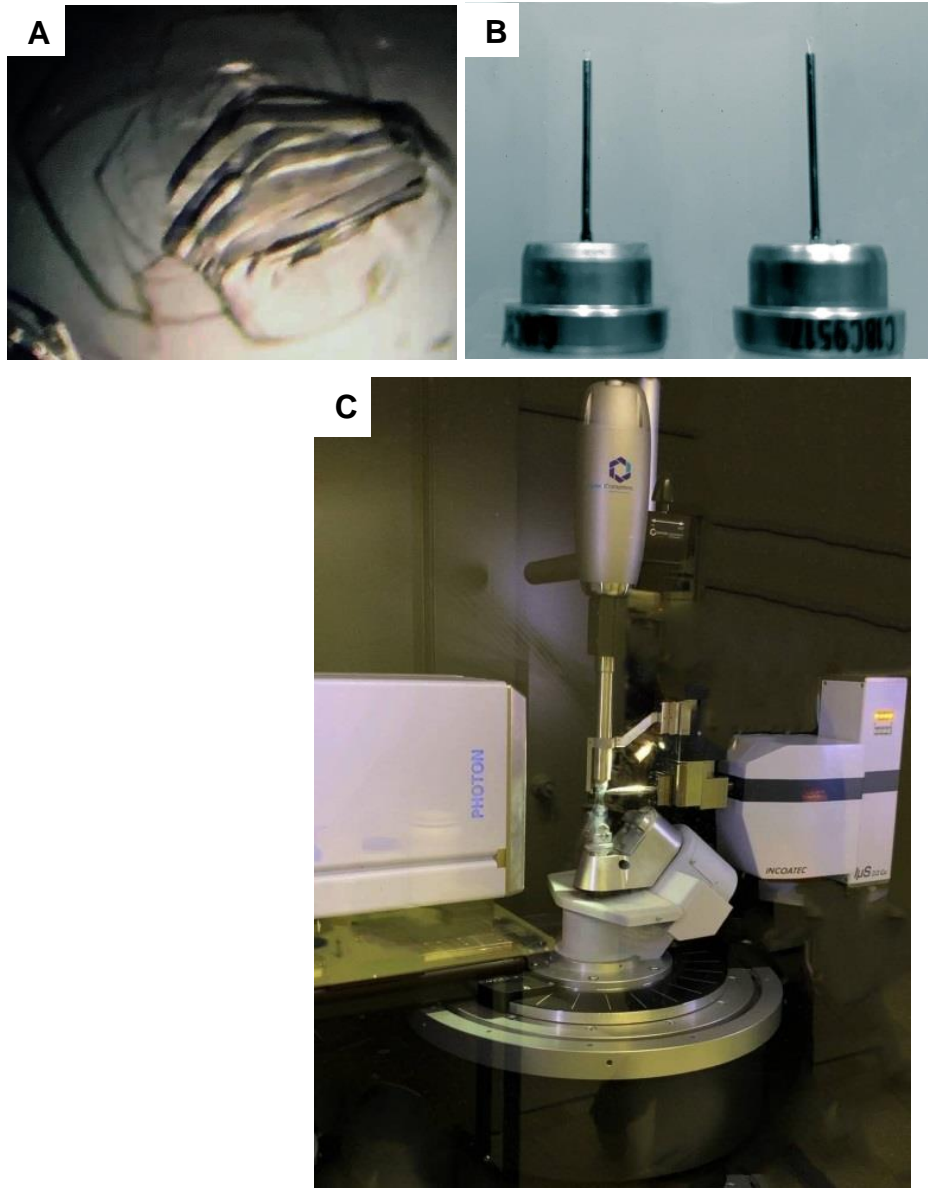


Figura 8. **A)** Cristal de NOG1. **B)** *Loops* para coleta de cristais. **C)** Difratorômetro D8 Venture localizado no Laboratório de Biofísica Molecular da UnB.

5 RESULTADOS E DISCUSSÃO

5.1 Expressão

Os genes das duas cistatinas de plantas selecionadas para este trabalho, CAF1 (12053 Da) e NOG1 (12414 Da), foram adquiridos da empresa GenOne já clonados no vetor pET24a, conforme descrito nos métodos. Os dois genes passaram pelas etapas de transformação em cepas de *E. coli*. De cada placa de cultura foi selecionada uma colônia para o pré-inóculo, que consistiu em adicionar esta colônia em um tubo Falcon de 15 mL com 5 mL de meio LB KLAN/CLO, deixados em overnigth (\cong 14 h a 37 °C ,sob agitação 220 rpm).

Levando em consideração a experiência prévia do laboratório em ensaios de expressão com o meio auto-indutor (ZYM-5052), os ensaios iniciais foram realizados neste meio e obtivemos bons resultados. Dessa forma, não foram realizados ensaios com outros meios de cultura.

Estes ensaios foram realizados com um mL do pré-inóculo e 500 mL de meio auto-indutor ZYM-5052 (Studier, 2005), por 10 h 45 min a 37 °C e 220 rpm, utilizando a cepa *E. coli* Lemo21 (D3), sendo realizados 3 litros de meio de cultura no final. O resultado neste meio foi igual para as duas proteínas, e, respectivamente, as Fig. **9A** e **9B** mostram como as PhyCys estavam expressas.

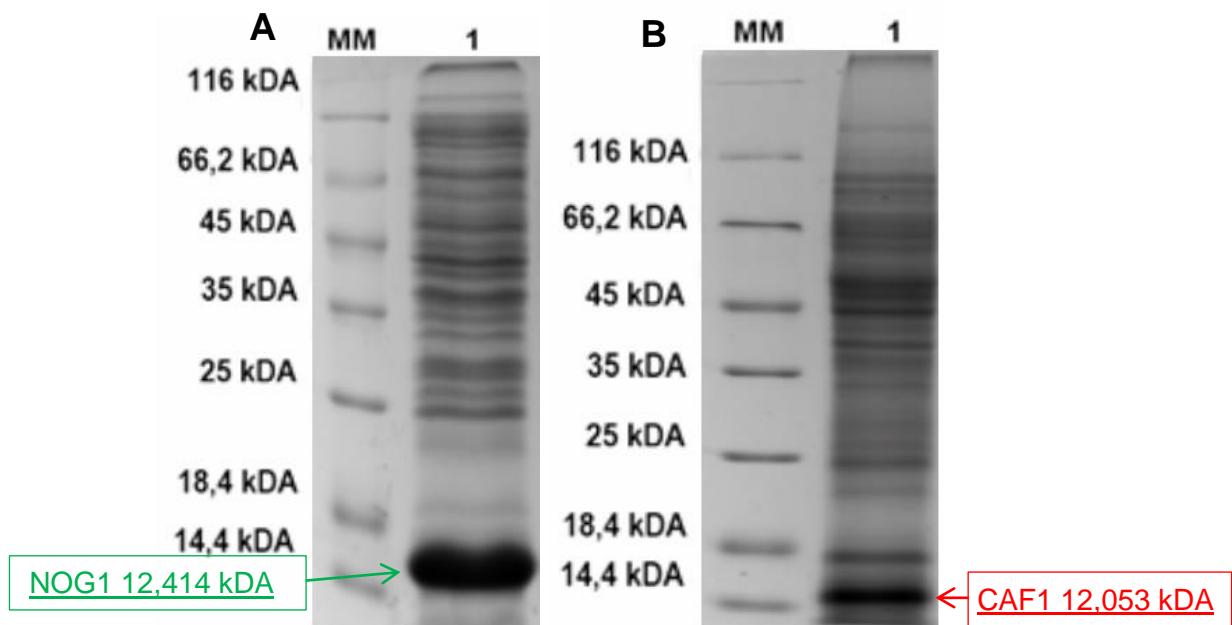


Figura 9. Eletroforese em SDS–PAGE (15 %) das cistatinas CAF1 e NOG1. Géis corados com Coomassie Brilliant Blue. Alíquota da expressão realizada e 3 litros de cultura em meio auto–indutor ZYM–5052 por 10h e 45 min. Representação dos poços nos géis: **A)** **MM** – marcador de peso molecular e **1** – Lisado bruto da NOG1 (12414 Da). **B)** **MM** – marcador de peso molecular e **1** – Lisado Bruto de CAF1 (12053 Da).

Obtivemos uma expressão otimizada das proteínas NOG1 e CAF1 em meio auto–indutor, este resultado está de acordo com artigos que indicam que o meio auto–indutor pode ser uma alternativa para um maior rendimento de proteínas quando comparado a outros meios de cultura para expressão, como o LB (Berrow *et al.*, 2006; Busso *et al.*, 2008, Moura, 2019).

5.2 Purificação

As purificações das proteínas foram realizadas por dois tipos de cromatografias. Primeiramente, utilizou–se a cromatografia de afinidade por íons metálicos imobilizados (IMAC) para a purificação. Posteriormente as proteínas obtidas passaram pela segunda etapa de purificação em cromatografia de exclusão molecular (SEC).

A cromatografia de afinidade está bem estabelecida para a purificação de proteínas, sendo considerada economicamente viável por recuperar altas produções e, também, permitir um alto grau de pureza destas proteínas (Pina *et al.*, 2014).

As construções de PhyCys utilizadas neste trabalho, por possuírem uma 6xHis tag, possibilitam a utilização de uma coluna de afinidade por níquel. Em cada IMAC realizada foram injetados 250 mL do sobrenadante da lise suplementado com 15 mM de imidazol, pois o imidazol proporciona a remoção de alguns contaminantes que interajam fracamente com a resina da coluna. Durante a purificação foram realizadas lavagens com um tampão de eluição, em diferentes concentrações (50 mM e 500 mM) de imidazol. Os 250 mL inicialmente injetados foram obtidos a partir de 3 litros da cultura de expressão. Este volume de cultura foi padronizado para CAF1 e NOG1, pois proporcionou a obtenção de grandes quantidades de proteínas solúveis, como é possível observar nas intensidades da banda dos poços dos géis de CAF1 e NOG1 (**Fig. 10**).

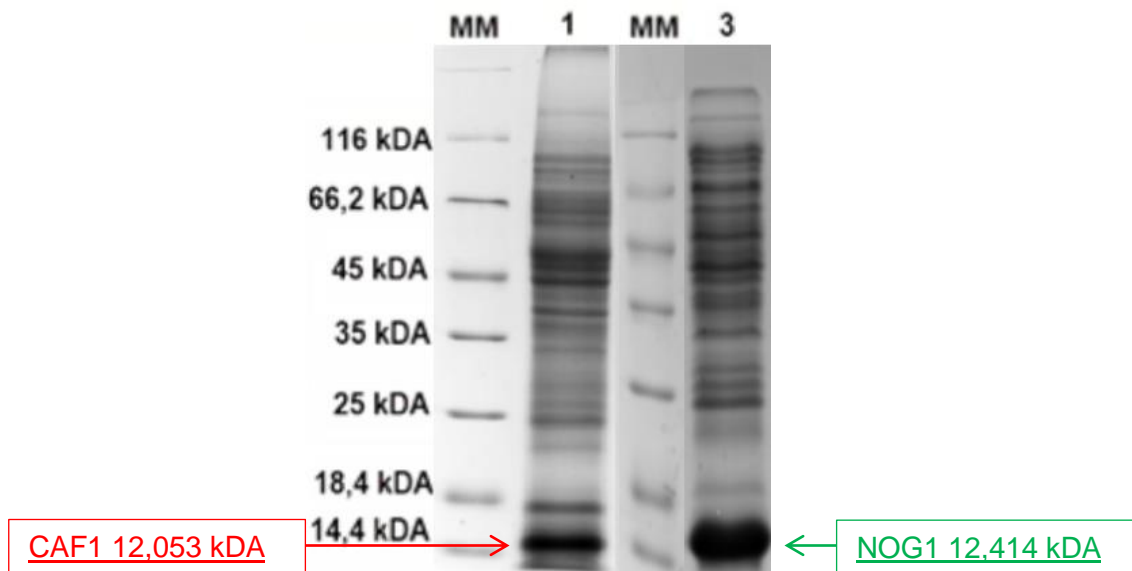


Figura 10. Eletroforese em SDS–PAGE 15% do sobrenadante da lise. Foram injetados nas colunas *HisTrap* HP 250 mL dos sobrenadantes das lise, estes volumes foram obtidos a partir de 3 litros de meio de cultura. O peso molecular da CAF1 é 12053 Da e o da NOG1 é 12414 Da. **MM** – Marcador de peso molecular; **1** – Sobrenadante da lise de CAF1, faixa intensa em 14,4 kDa de peso molecular CAF1, mostrando bastante proteína solúvel; **3** – Sobrenadante da lise de NOG1, faixa bastante intensa perto de 14,4 kDa, também mostrando proteína abundante.

Durante as IMAC realizadas foi coletada a fração que atravessa a coluna livremente (*flowthru*), uma vez que quando se injetam as amostras, uma parte destas proteínas não se liga à coluna, ficando na fração não adsorvida, como pode ser observado nos géis, tanto de CAF1 (**Fig. 11**), quanto de NOG1 (**Fig. 12**).

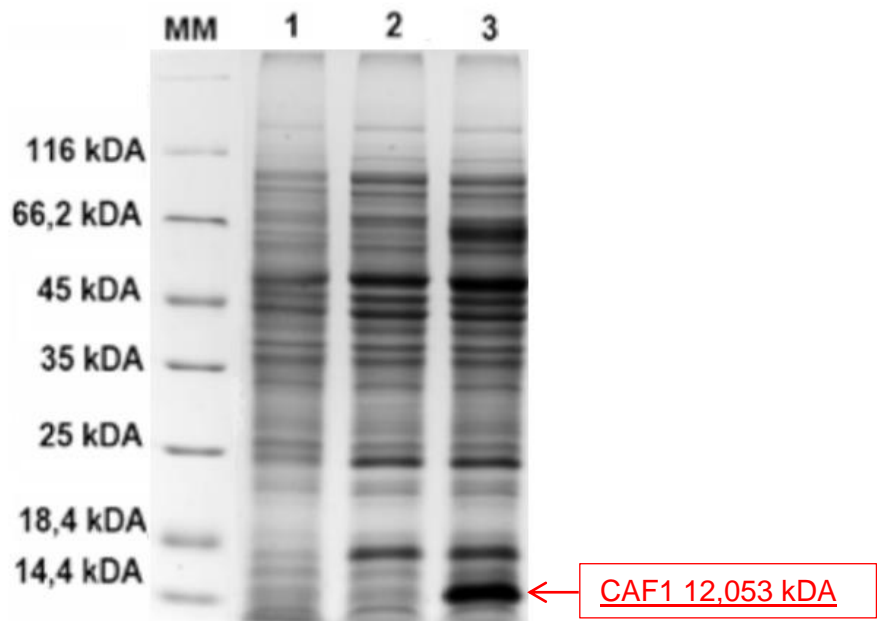


Figura 11. Uma parte da proteína atravessava a coluna e não se ligava a ela. Desta forma, foram retiradas alíquotas do *flowthru* durante o procedimento de cromatografia de afinidade e com estas alíquotas realizou-se a eletroforese em SDS-PAGE (15%) de diferentes eluições do *flowthru* de CAF1(12053 Da), para padronização de coleta do *flowthru*. **MM** – Marcador de peso molecular; **1** – *Flowthru* eluído em 50 mL; **2** – *Flowthru* eluído em 100 mL; **3** – *Flowthru* eluído em 150 mL.

Nas IMAC de CAF1, e, de cada *flowthru* coletado foi possível observar que ao passar 150 mL da proteína pela coluna, a mesma não se ligava mais e saía em abundância.

Na primeira IMAC da NOG1, o *flowthru* foi coletado em duas etapas (**Fig. 12**). Por análise eletroforética do gel SDS-PAGE, deduz-se, pela intensidade da banda, que a quantidade de proteína da IMAC do *flowthru* eluído a partir de 150 mL era abundante, portanto padronizamos apenas coletar esta fração.

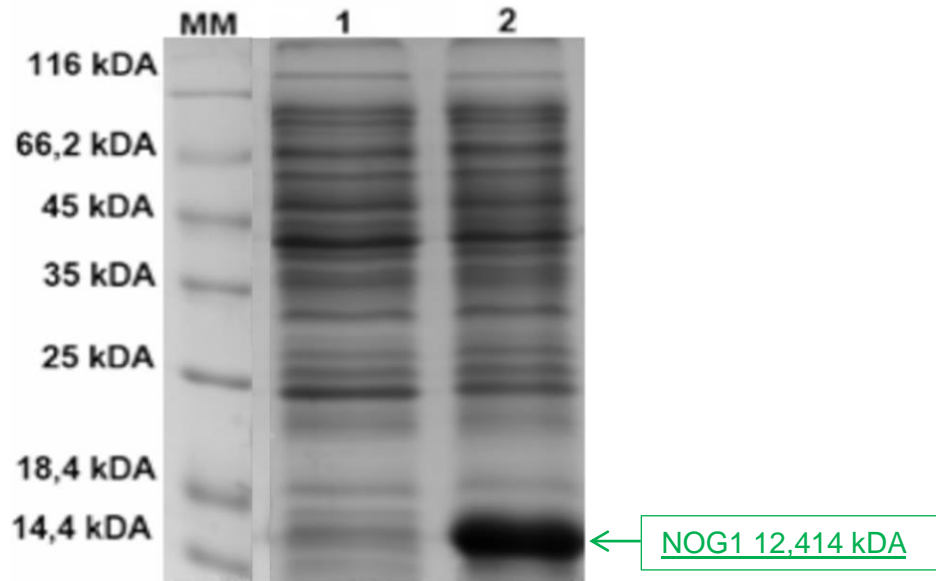


Figura 12. Parte da proteína não se liga à coluna. Desta forma, ela atravessa livremente (*flowthru*), portanto foram retiradas alíquotas do durante o procedimento de cromatografia de afinidade e com estas alíquotas realizou-se a eletroforese em SDS-PAGE (15 %) de diferentes eluições do *flowthru* de NOG1 (12414 Da), para padronizar em qual momento se coletaria o *flowthru* **MM** – Marcador de peso molecular; **1** – *Flowthru* eluído em 50 mL; **2** – *Flowthru* eluído em 150 mL.

Na primeira IMAC realizada para purificação da NOG1 foi obtida grande quantidade da proteína (**Fig. 13**). Já na re-injeção do *flowthru* coletado na primeira purificação observamos apenas um pico, que supomos se tratar da proteína predominantemente monomérica (**Fig. 14**). De cada etapa se retirou uma alíquota para SDS-PAGE, onde foi observada grande quantidade de proteína NOG1 (12414 Da) (**Fig. 15**), e, também, foram observados no poço 2 e 3 da **Figura 15** bandas de alto peso molecular, por volta de 60 kDa e 80 kDa, que aparentemente se tornam mais intensas na medida em que a banda referente a NOG1 se torna intensa. Todas as amostras foram coletadas e sempre foram armazenadas em gelo.

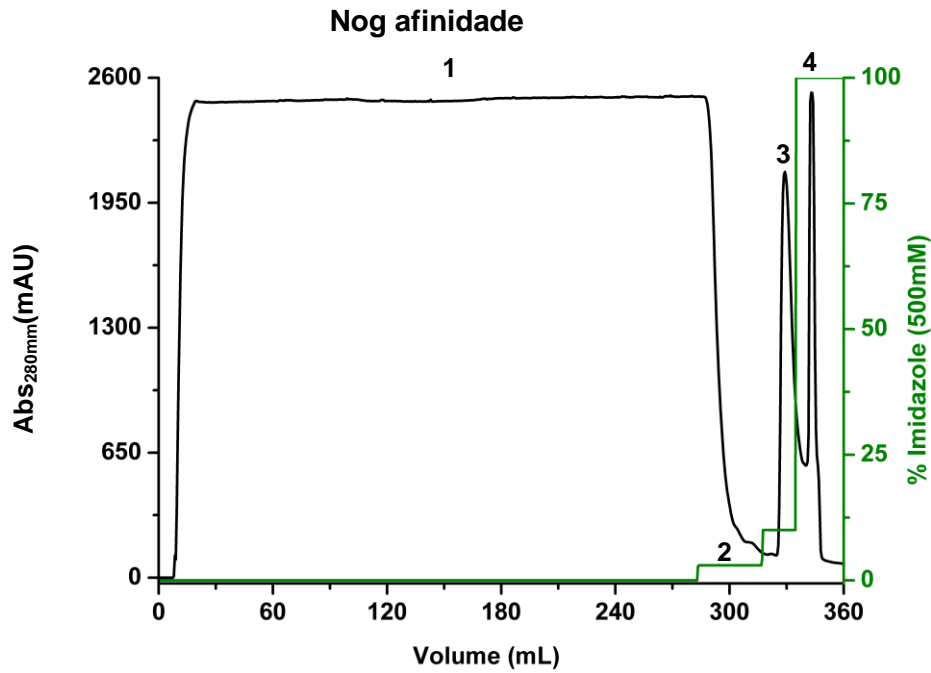


Figura 13. Cromatograma de purificação por afinidade a níquel, injeção de 250 ml da proteína NOG1. **1** – *Flowthru*; **2** – Lavagem com 15 mM de imidazol; **3** – Eluição da proteína com 50 mM de imidazol. **4** – Eluição da proteína com 500 mM de imidazol.

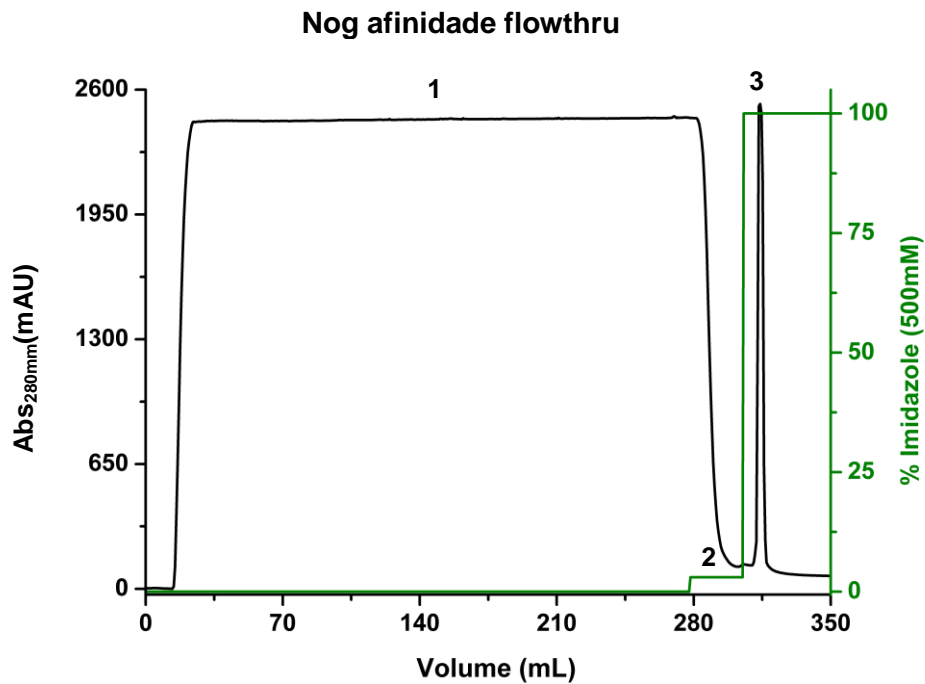


Figura 14. Injeção do *flowthru* coletado previamente da proteína NOG1 (recromatografia de afinidade). **1** – *Flowthru*; **2** – Lavagem com 50 mM de imidazol; **3** – Eluição da proteína com 500 mM de imidazol.

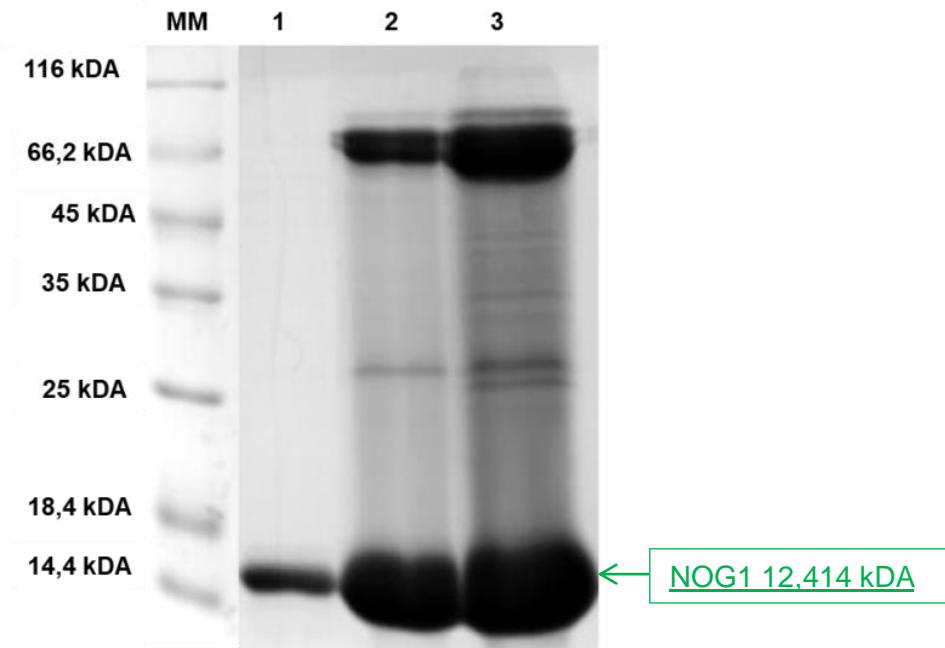


Figura 15. Eletroforese em SDS–PAGE (15%) das amostras de purificação de NOG1 (12414 Da), amostras da cromatografia de afinidade realizada com 250 mL de sobrenadante da lise. Foram realizadas diferentes lavagens da coluna e as respectivas alíquotas estão apresentadas no gel. **MM** – Marcador de peso molecular; **1** – NOG1 eluída em 15 mM de imidazol; **2** – NOG1 eluída em 50 mM de imidazol; **3** – NOG1 eluída em 500 mM de imidazol. É possível observar que mesmo em diferentes lavagens se nota grande quantidade da proteína.

As amostras obtidas nas etapas de cromatografia por afinidade foram utilizadas em procedimentos de cromatografia por exclusão molecular (SEC) em coluna Superdex 75 (10/300). Essa etapa visa aumentar a pureza das amostras e tentar separar diferentes estados oligoméricos das cistatinas, principalmente obter dímeros e monômeros puros. Valadares e colaboradores (2013) já haviam apresentado dados que a cromatografia por afinidade a níquel permitia isolar uma fitocistatina, produzida por expressão heteróloga, porém ao realizar a SEC foi observado mais de um pico. A resolução da estrutura cristalográfica aliada a dados de ressonância magnética nuclear demonstraram a formação de dímeros por troca de domínios. Júnior e colaboradores (2016) também encontraram resultados semelhantes para outra fitocistatina.

Nas SEC das amostras coletadas na primeira IMAC de NOG1 (obtidas na **Fig. 13**), constata-se que os picos de dímeros e tetrâmeros apresentam maior

absorbância que os picos relativos a monômeros. Já na amostra referente ao pico eluído com 500 mM de imidazol na **Figura 14**, foi observado um pico muito intenso e largo, no volume de eluição referente a monômeros (pico cor de rosa na **Fig. 16**).

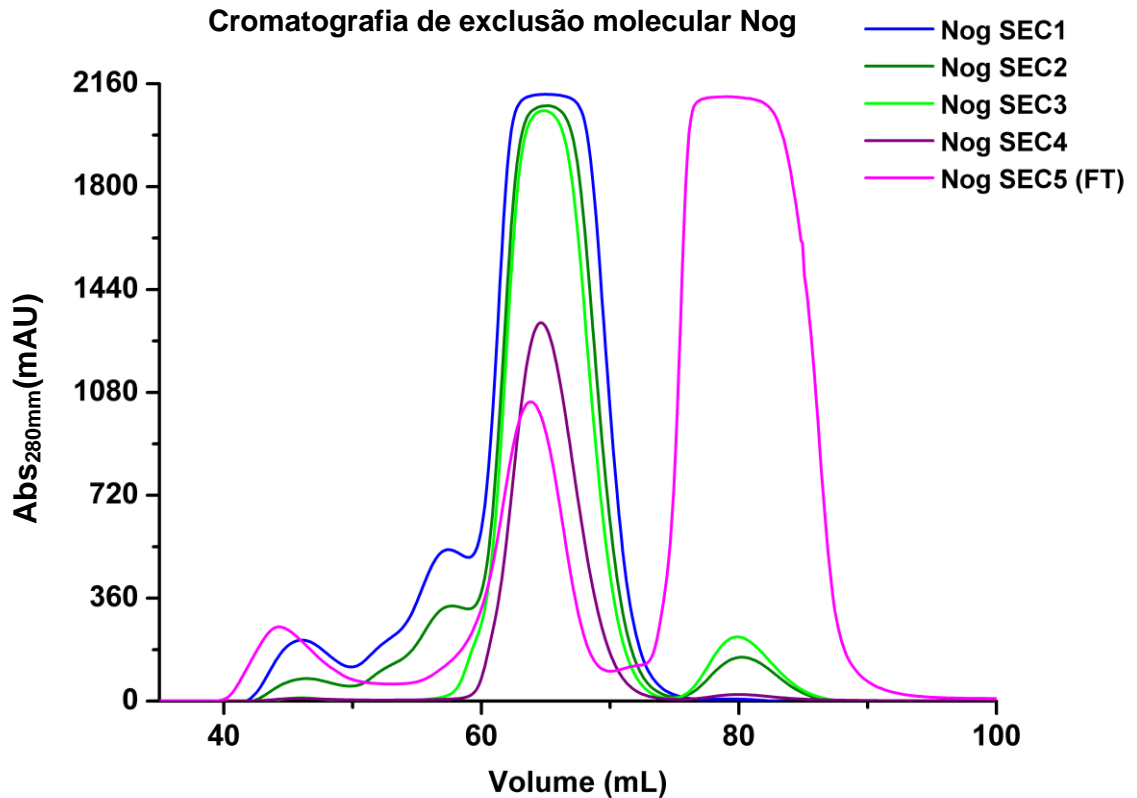


Figura 16. Sobreposição de cromatogramas das diferentes injeções na SEC, das amostras obtidas da purificação de NOG1 nas IMAC. Cada curva representa uma corrida da SEC. **SEC1** – Cor azul, bastante oligômeros e dímeros. **SEC2** – Cor verde escuro, um pouco de oligômeros, bastante dímeros e um pouco de monômeros; **SEC3** – Cor verde claro apresenta bastante dímeros e um pouco de monômeros; **SEC4** – Cor roxa, bastante dímeros e **SEC5 (FT)** – Cor de rosa, está é a injeção do *flowthru*, e apresenta predominantemente monômeros, mas também possui dímeros.

A **Figura 17** deixa claro que amostras provenientes de um primeiro procedimento de purificação por IMAC apresentam predominantemente dímeros, e que ao reutilizarmos o *flowthru* obtido nessa primeira purificação em um segundo procedimento de IMAC obtemos majoritariamente monômeros. A figura apresenta dois cromatogramas e alíquotas correspondentes a diferentes volumes de eluição. No gel de SDS–PAGE foram utilizadas cada alíquota. Amostras eluídas por volta de 65 mL correspondem a dímeros e amostras eluídas em torno de 80 mL são monoméricas. O SDS–PAGE deixa claro que se trata da mesma proteína.

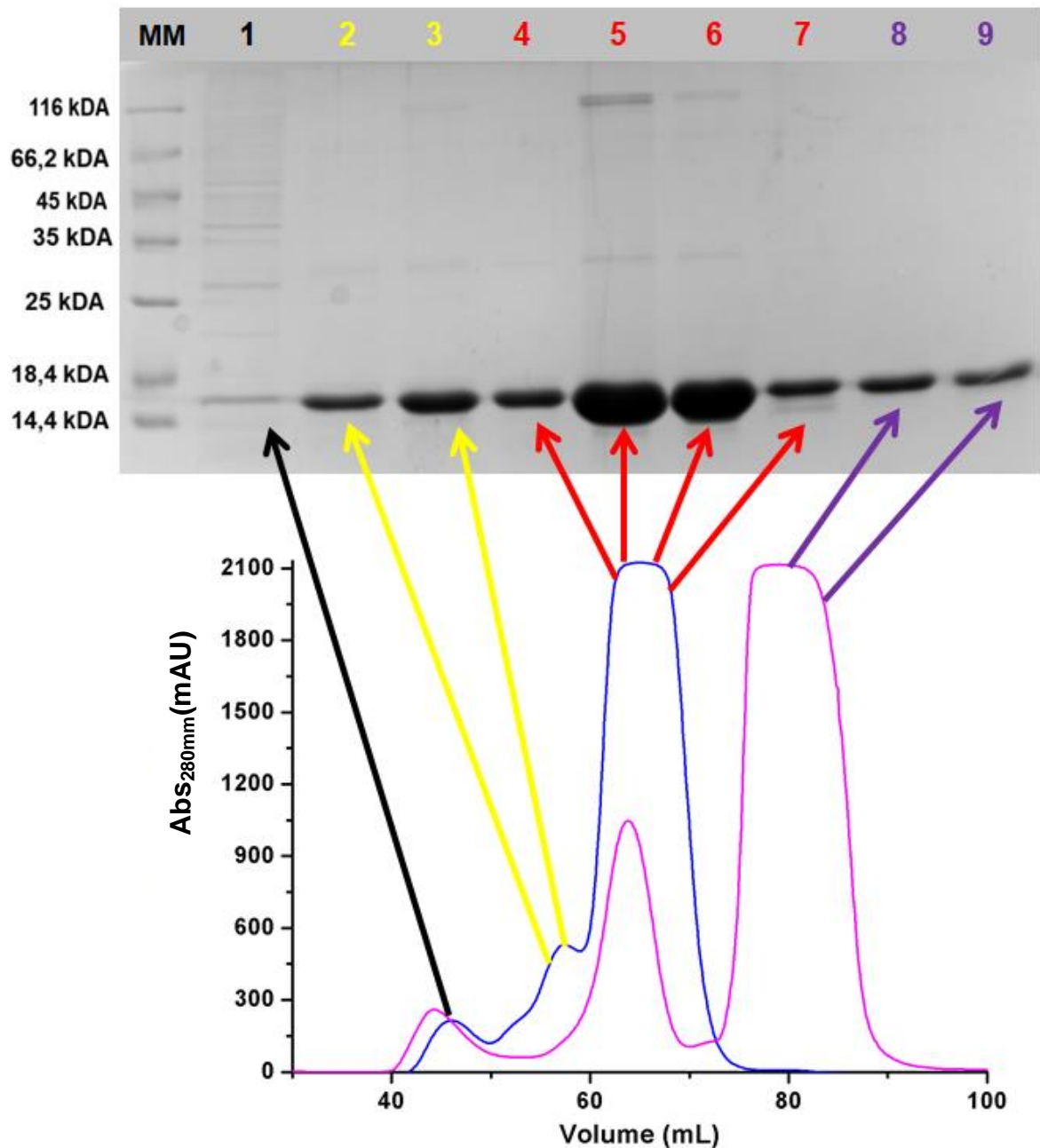


Figura 17. Sobreposição de dois cromatogramas de exclusão molecular da proteína NOG1 (12414 Da). Amostras provenientes de diferentes IMAC foram injetadas em uma coluna Superdex 75, sendo primeira amostra obtida na IMAC do sobrenadante da lise (azul) e a segunda corresponde a IMAC da injeção do *flowthru* (rosa). Ou seja, a curva rosa corresponde ao sobrenadante após a primeira passagem pela coluna, quando ficaram retidas as proteínas que quando eluídas deram origem à curva azul. Eletroforese em SDS-PAGE das amostras eluídas nas purificações por SEC. As setas indicam as diferentes amostras coletadas. **MM** – marcador de peso molecular; **1** – Void da SEC (43 mL); **2** – 54 mL; **3** – 57 mL; **4** – 60/61 mL; **5** – 63/65 mL; **6** – 67/69 mL; **7** – 71/72 mL; **8** – 79/81 mL; **9** – 81/83 mL.

Nas purificações da CAF1 por afinidade percebemos que ao grande rendimento, apesar de obtermos um grande volume de proteína pura na eluição de 500mM de imidazol (**Fig. 18**, alíquota do gel **Fig. 20** poço 4), ainda restava muita proteína de interesse no *flowthru*. Após esta observação decidiu-se pela reinjeção do *flowthru* obtido na primeira purificação em um novo procedimento de purificação por afinidade onde também foi obtido um excelente rendimento (**Fig. 19**). De cada uma das etapas da primeira IMAC (**Fig. 18**) retirou-se uma alíquota para SDS-PAGE que indica a presença de CAF1 em grande quantidade (**Fig. 20**).

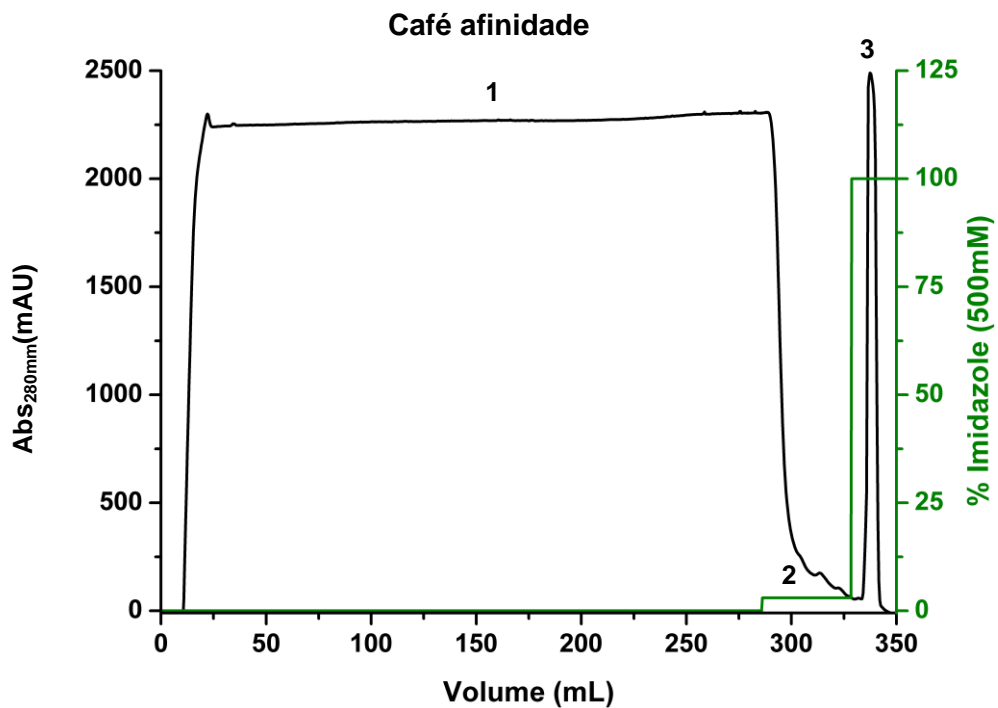


Figura 18. Cromatograma de purificação por afinidade a níquel, injeção de 250 mL da proteína CAF1. **1** – *Flowthru*; **2** – Lavagem com 15 mM de imidazol; **3** – Eluição da proteína com 500 mM de imidazol.

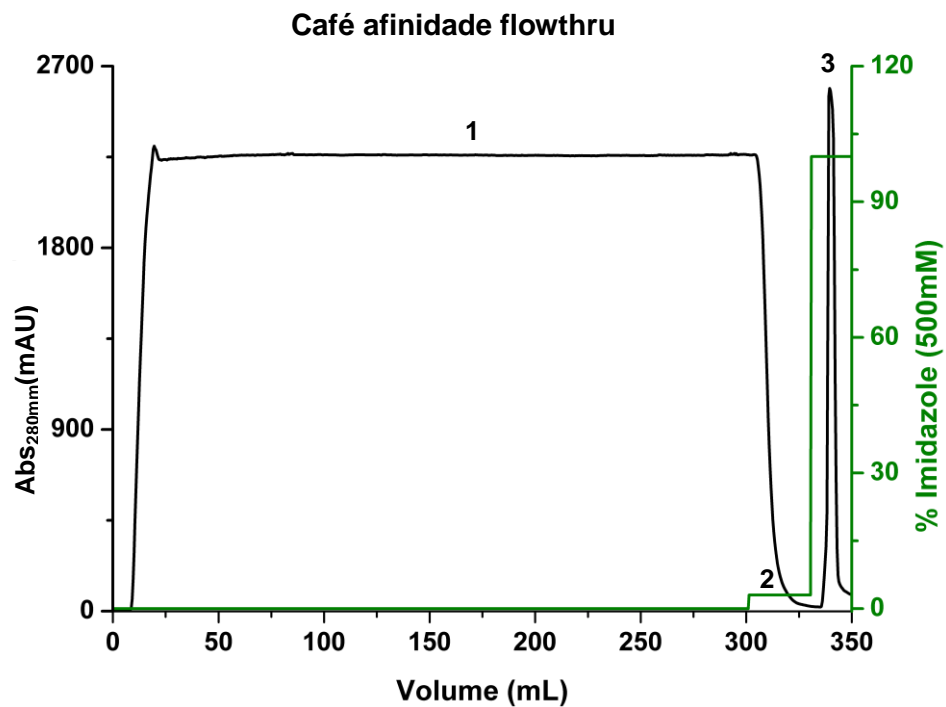


Figura 19. Cromatograma de purificação da proteína CAF1 por afinidade a níquel, injeção do *flowthru* coletado previamente no cromatograma apresentado na **Figura 18** (recromatografia de afinidade). **1** – *Flowthru*; **2** – Lavagem com 15 mM de imidazol; **3** – Eluição da proteína com 500 mM de imidazol.

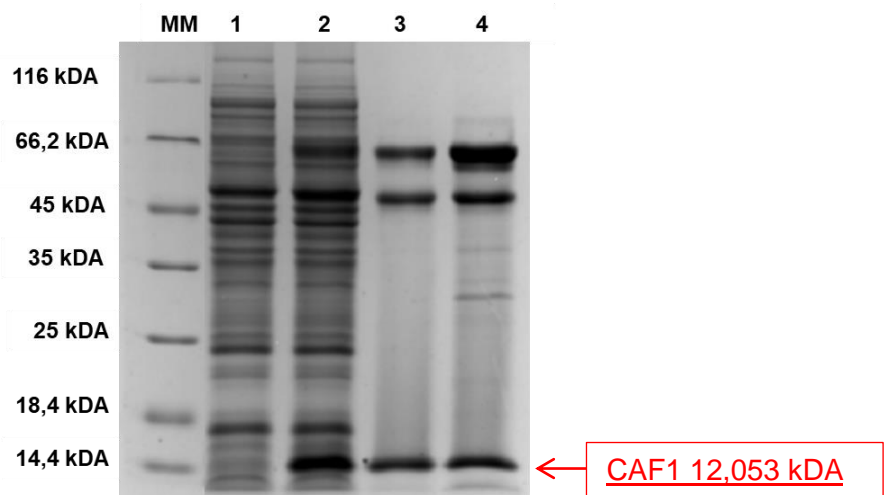


Figura 20. Eletroforese em SDS-PAGE (15 %) das amostras de purificação de CAF1 (12053 Da). Amostras de cromatografia de afinidade, a representação no gel dos poços 3 e 4 foram da IMAC realizada com 250 mL de sobrenadante da lise e os poços 1 e 2 são amostras coletadas do *flowthru*, em dois momentos diferentes. Foram feitas duas lavagens da coluna e as alíquotas retiradas delas estão apresentadas no gel. **MM** – Marcador de peso molecular; **1** – Amostra do início do *flowthru* (eluído em 50 mL) **2** – Amostra do *flowthru* eluída a partir de 150 mL; **3** – Amostra da IMAC, fração eluída em 15 mM de imidazol; **4** – Amostra de CAF1, da IMAC, fração eluída em 500 mM de imidazol. Nota-se no gel parte da proteína no *flowthru* (poço 2), e nos poços 3 e 4 é possível observar que mesmo com diferentes lavagens se obteve bastante proteína.

Com as amostras obtidas nas IMAC da CAF1 foram realizadas as SECs com uma coluna Superdex 75 (10/300) para obtenção de proteínas puras, pois todos os géis apresentaram uma banda dupla em alto peso molecular, sugerindo a existência de outras formas oligoméricas da proteína ou contaminantes.

A **Figura 21** ilustra cromatogramas de SEC de amostras de CAF1. Nas SEC de amostras primárias de IMAC da CAF1 (curva azul), constata-se que o pico relativo a monômeros é bem maior que o de dímeros e tetrâmeros, ao contrário das SEC de NOG1, contudo CAF1 geralmente apresenta estes outros dois picos. Já na SEC da injeção do *flowthru*, a maior parte da proteína é eluída no pico referente a monômeros (**Fig. 21**, curva vermelha).

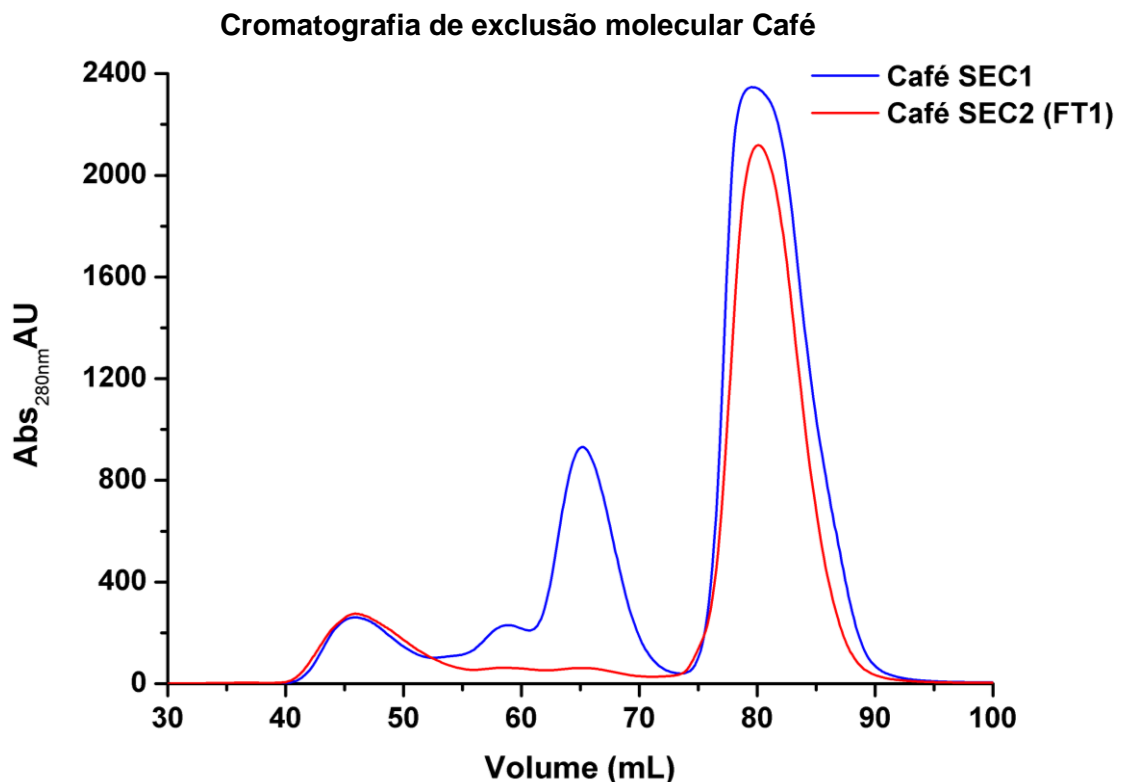


Figura 21. Sobreposição de cromatogramas de SEC oriundos da injeção das amostras obtidas nas purificações por IMAC de CAF1. Cada cor representa um procedimento de SEC, a curva em vermelho representa a amostra correspondente a IMAC do *flowthru*, apresentado na **Figura 19**.

Das amostras de CAF1 na SEC1 (café SEC1) foram retiradas alíquotas para SDS-PAGE, pois a presença de mais de um pico na curva azul da **Figura 21**

indica a existência de diferentes proteínas ou diferentes estados oligoméricos da CAF1. No gel apresentado na **Figura 22**, cada poço representa uma fração coletada em determinado volume da cromatografia. Com exceção do *void* da SEC, todas as amostras coletadas apresentam uma faixa de alta intensidade com peso molecular correspondente à proteína CAF1. Isso indica que a proteína obtida em diferentes picos na SEC corresponde a CAF1 em diferentes estados oligoméricos. Em todos os poços também foram observadas duas bandas por volta de 50 kDa e 66 kDa que aparentemente se tornam mais intensas na medida em que a banda referente a CAF1 se torna intensa. Não sabemos explicar essa banda, pois o SDS deveria desnaturar as proteínas no GEL, e no procedimento de SEC uma proteína de 66 kDa deveria ser eluída em um volume bem menor que o observado no poço 8 da **Figura 22** (82/84 mL). Todo esse procedimento foi repetido algumas vezes e essas bandas de alto peso molecular que eluem em altos volumes na SEC sempre se mantiveram presentes.

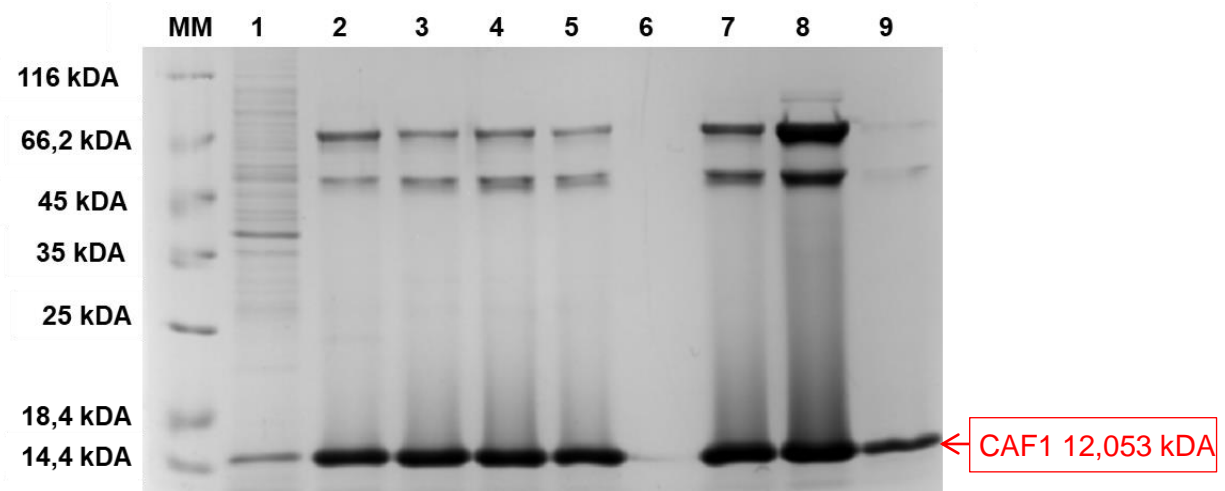


Figura 22. Eletroforese em SDS–PAGE das amostras obtidas na purificação da CAF1, na SEC, sendo cada poço a amostra coletada em um volume de eluição. **MM** – marcador de peso molecular; **1** – *Void* da SEC; **2** – 58/59 mL; **3** – 63/64 mL; **4** – 64/66 mL; **5** – 66/68 mL; **6** – Vazio; **7** – 76/78 mL; **8** – 82/84 mL; **9** – 88/89 mL. Os diferentes volumes representam tetrâmeros, dímeros e monômeros, respectivamente.

Há muito tempo se sabe que cromatografia de exclusão por tamanho (SEC) é uma ferramenta eficiente para separar moléculas por tamanho (Ricker e Sandoval, 1996; Bevan e Stretton, 2016). Obtivemos monômeros, dímeros e tetrâmeros no procedimento da SEC. A SEC é uma técnica de separação não destrutiva que pode separar homoproteínas, elevando o grau de pureza para a cristalização.

5.2.1 Calibração da coluna Superdex75 (10/300)

Para estimar a massa molecular aparente das proteínas NOG1 e CAF1 utilizando a técnica de cromatografia de exclusão molecular (SEC), foi necessária a calibração da coluna Superdex75 (10/300), produzida pela empresa da GE e utilizada em todas as purificações e ensaios de SEC aqui apresentados. Para tanto foi utilizado o *Gel Filtration LMW Calibration Kit*, obtido da empresa GE.

A calibração foi feita de acordo com as recomendações do manual do kit de calibração. Para tanto, misturas das proteínas padrão foram preparadas e injetadas na coluna, e o *void* da coluna foi determinado utilizando o corante *Dextran Blue 2000*. Para a calibração foi utilizado o mesmo tampão que foi utilizado nos ensaios de oligomerização (20 mM tris pH 7,5, 150 mM NaCl). O resultado do processo de calibração está apresentado na **Figura 23**.

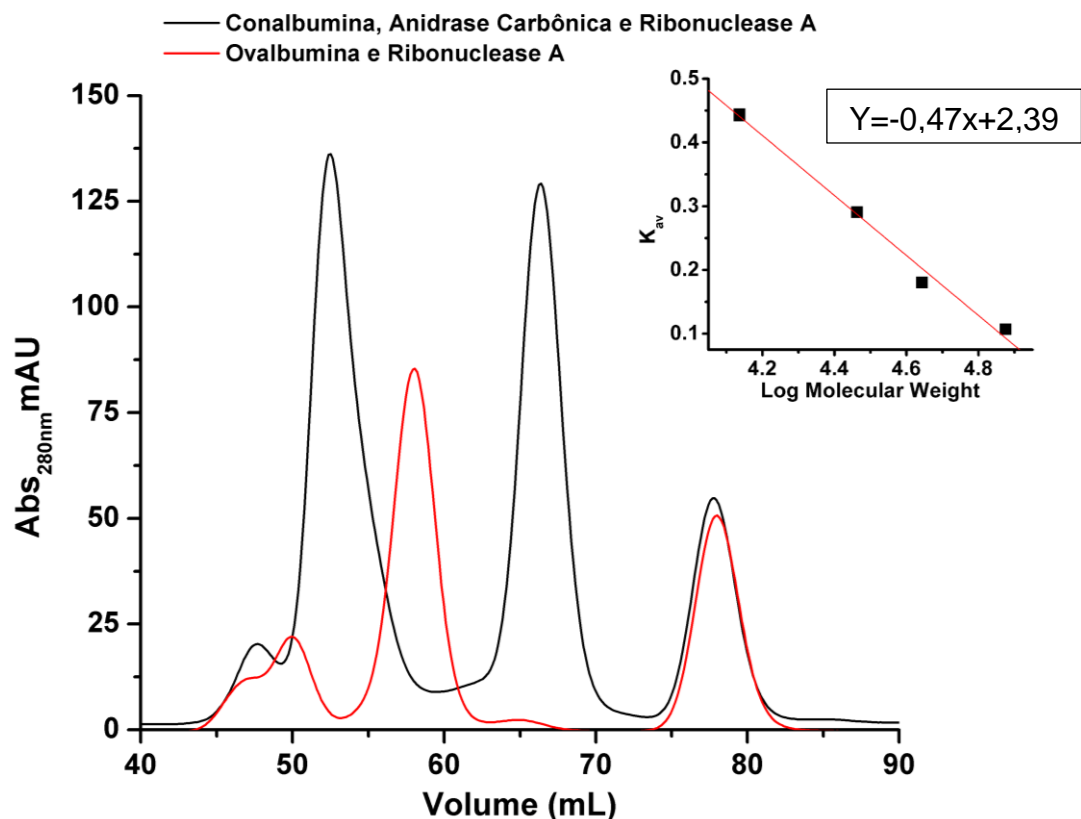


Figura 23. Cromatograma resultante do processo de calibração da coluna Superdex 75 (10/300). A curva preta foi obtida pela injeção da mistura das proteínas conalbumina, anidrase carbônica e ribonuclease A. A curva vermelha foi obtida pela injeção das proteínas ovalbumina e ribonuclease A. O gráfico inserido na parte superior direita mostra o ajuste obtido utilizando os valores dos picos produzidos pelas proteínas injetadas.

É possível utilizar o ajuste obtido para estimar a massa molecular de proteínas eluídas na coluna calibrada. A proteína NOG1 possui massa molecular estimada pela sequência 12,4 kDa, e a proteína CAF1 possui massa molecular estimada pela sequência 12,1 kDa. Ambas, geralmente, podem ser eluídas em dois picos, um atribuído a monômeros em torno de 79 mL, o que corresponde a uma massa estimada de 12,6 kDa, e outro correspondente a dímeros por volta de 64 mL, e de acordo com a calibração possui massa estimada de 33,3 kDa. Esses são valores aproximados, a análise de todos os cromatogramas de SEC apresentados nessa dissertação mostra que os valores estimados para o pico de monômeros de NOG1 variam entre 12,3 kDa e 12,6 kDa, ao passo que os valores para o pico de dímero variam entre 33,4 kDa e 34,2 kDa. Já para CAF1 são observados os intervalos de 11,2 kDa a 12,6 kDa para monômeros, e 30,7 kDa a 33,3 kDa para a massa estimada de dímeros.

Considerando os valores apresentados no parágrafo anterior, a massa molecular estimada para os monômeros de NOG1 e CAF1 estão de acordo com os valores esperados. Contudo, dímeros de NOG1 e CAF1 possuem massa 24,8 kDa e 24,2 kDa, respectivamente, e os valores de massa estimados para essas espécies se encontram acima do esperado. Essa discrepância já era antecipada, pois em contraste com a estrutura aproximadamente globular do monômero, o dímero observado na estrutura da NOG1 apresenta uma estrutura alongada, com dois lóbulos conectados por duas longas fitas beta antiparalelas. Adicionalmente, a literatura descreve casos semelhantes, por exemplo, o valor de massa molecular da canacystatin-1 estimada pela sua sequência é de 11,9 kDa (Valadares, 2013). No entanto, a canacystatin-1 foi eluída em 3 picos distintos na SEC, com massas estimadas em 14,5 kDa, 33,4 kDa e 59,3 kDa. Assim, é esperado que a estimativa de massa pela SEC de dímeros por troca de domínios de fitocistatinas tenham sua massa superestimada.

5.3 Caracterização do Perfil de Oligomerização

As PhyCys CAF1 e NOG1 passaram por ensaios para a caracterização das populações em solução onde alguns parâmetros foram inicialmente testados e estabelecidos.

A primeira informação importante obtida por ensaios utilizando SEC foi que amostras de dímeros ou monômeros previamente isoladas por SEC não sofrem interconversão significativa se mantidas em gelo, mesmo em concentrações de até 10 mg/mL e após longos períodos, como duas semanas. A identificação dessa característica dessas fitocistatinas permitiu o delineamento e execução dos experimentos subsequentes.

Para verificar a influência da temperatura na interconversão de estados oligoméricos, inicialmente avaliamos quanto tempo as proteínas deveriam ser incubadas a alta temperatura (70 °C) para que o equilíbrio fosse alcançado. As proteínas foram aquecidas em banho–maria por diferentes tempos de aquecimento (de 1 a 40 min) e subsequentemente injetadas em uma coluna acoplada Superdex 75 (10/300) acoplada ao AKTA Prime. Foi observado que 20 min é um intervalo de tempo mais que suficiente para atingir o equilíbrio (**Fig. 24**). Este ensaio foi realizado utilizando dímeros de NOG1 (0,40 mg/mL), que foram aquecidos a 70 °C em banho–maria, com auxílio de um bloco térmico, por intervalos de tempo variáveis. Após o aquecimento, as amostras foram resfriadas e injetadas em uma coluna Superdex 75 (**Fig. 24**). Utilizando dessa informação, foi definido que todos os outros ensaios subsequentes fossem executados aquecendo as amostras por 20 min.

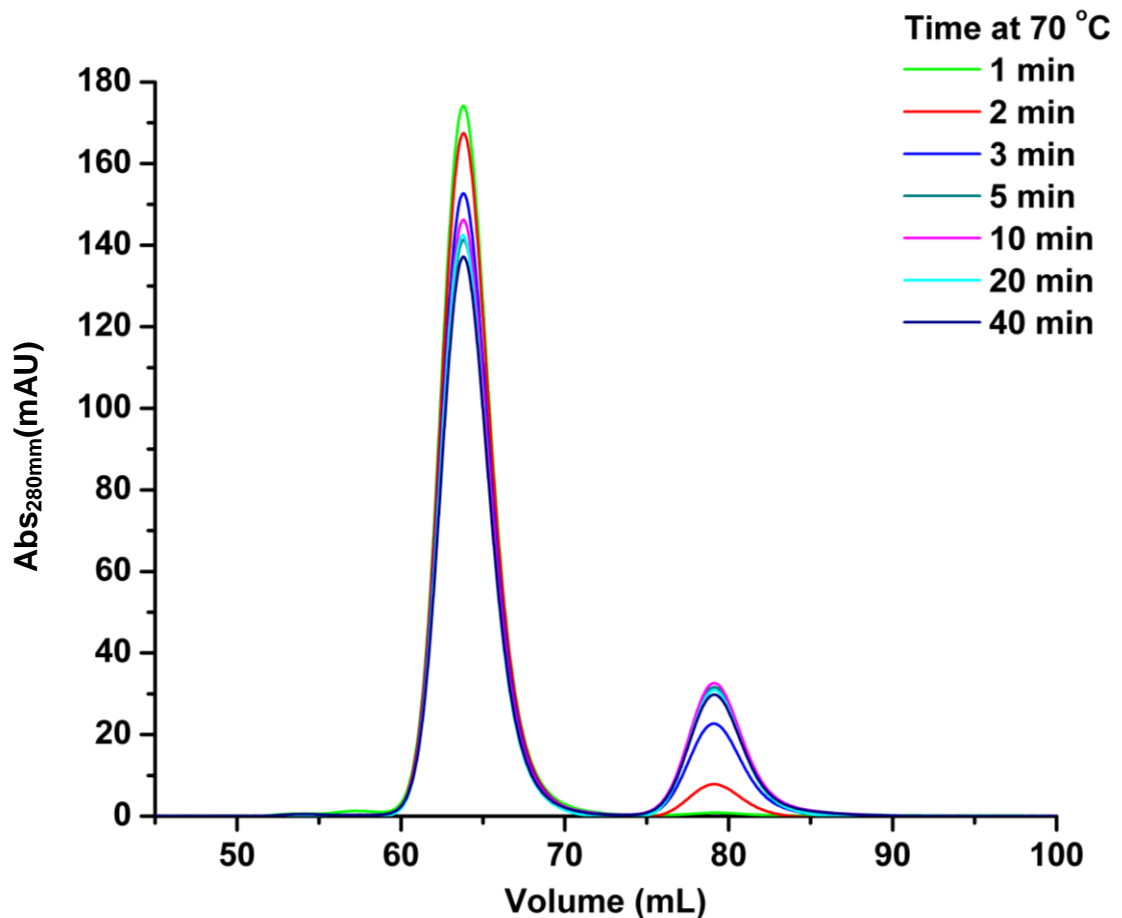


Figura 24. Gráfico de estabilidade de dímeros de NOG1 a 70 °C em função do tempo. O gráfico foi obtido pela sobreposição de sete cromatogramas de SEC de amostras idênticas de NOG1 aquecidas a 70 °C por intervalos de tempo distintos. O pico referente a dímeros ocorre em torno de 64 mL e o pico referente a monômeros está em torno de 79 mL.

Para verificar o efeito da temperatura na interconversão entre dímeros e monômeros, foi executado um ensaio aquecendo dímeros de NOG1 (0,40 mg/mL) em banho-maria por 20 min a diferentes temperaturas. Após o resfriamento e injeção dessas amostras na SEC, foi observado que os dímeros são bem estáveis e apenas uma parcela da população dimérica se transforma em população monomérica sob aquecimento. Analisando o cromatograma (**Fig. 25**), os dímeros aquecidos em: 40 °C quase não se transformavam em monômeros; 50 °C é observado um primeiro indício da presença de monômeros; 60 °C já se observa em uma população monomérica definida; 70 °C o pico referente aos monômeros atinge seu máximo; 80 °C o pico referente aos monômeros é ligeiramente menor que o observado a 70 °C, contudo também é observado um expressivo pico em torno de 44 mL, que corresponde ao *void* da coluna, onde espécies possuindo mais de 80

kDa são eluídas. Esse pico possivelmente corresponde a agregados, indicando que nessa temperatura a proteína possivelmente se desnaturou irreversivelmente formando agregados. Devido a essa perda de massa da amostra o pico em torno de 64 mL, referente aos dímeros, sofreu uma expressiva redução.

Em conjunto, esses resultados sugerem que amostras contendo dímeros da proteína NOG1 são relativamente estáveis por 20 minutos a 40 °C, pois nessas condições não foram observados monômeros ou agregados. Essas amostras também podem ser aquecidas a 70 °C por pelo menos 20 minutos sem que a formação de agregados seja observada por SEC. A temperatura de 70 °C também propiciou a maior dissociação de dímeros em monômeros.

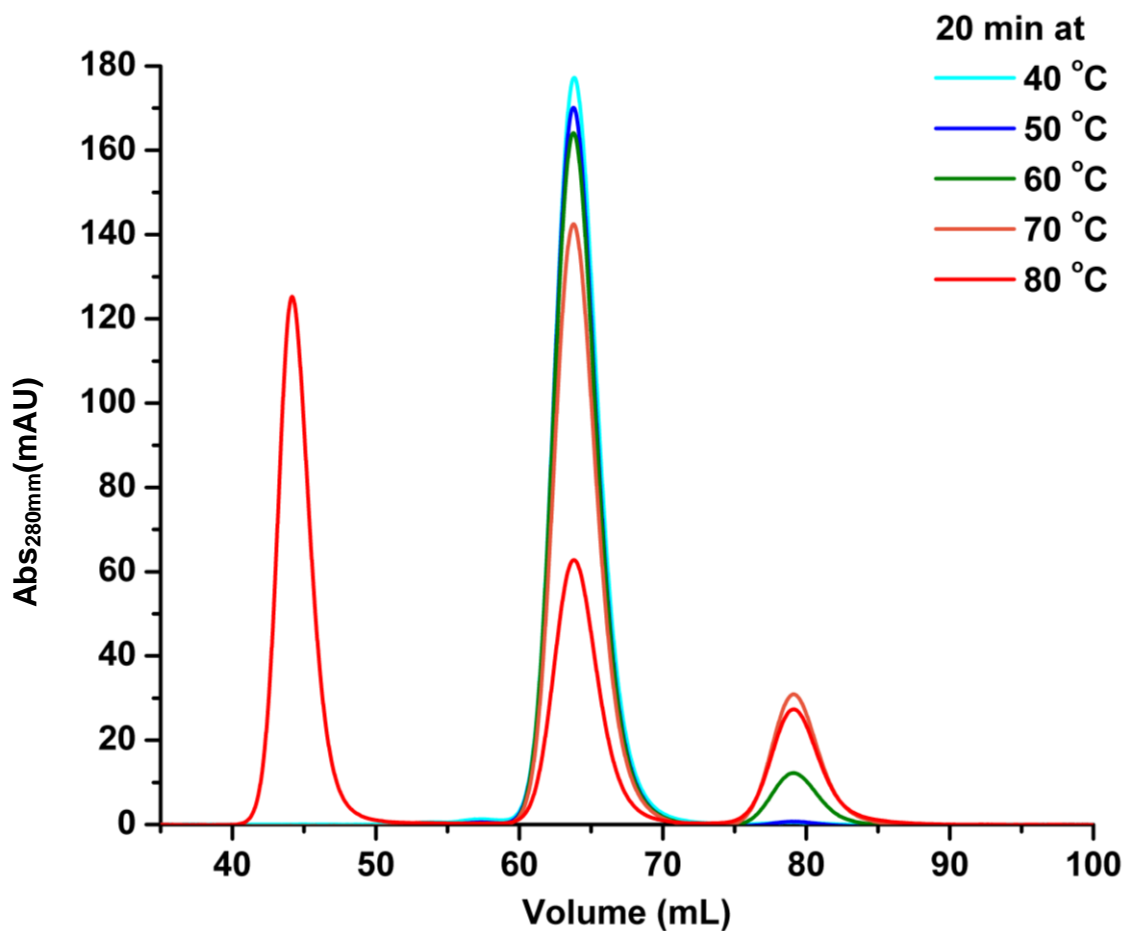


Figura 25. Gráfico da influência da temperatura na formação de monômeros a partir de dímeros de NOG1. O gráfico apresenta a sobreposição de cromatogramas obtidos a diferentes temperaturas. O pico referente a dímeros ocorre em torno de 64 mL e o pico referente a monômeros está em torno de 79 mL. O pico em torno de 44 mL corresponde ao *void* da coluna e representa estados oligoméricos de massa superior a 80 kDa. A curva vermelha (80 °C) apresenta a formação de oligômeros de alto peso molecular, correspondente ao pico eluído entre 40 e 50 mL.

De forma similar ao observado para NOG1, ao aquecer dímeros de CAF1 (0,16 mg/mL) a diferentes temperaturas, por 20 min em banho-maria, e em seguida resfriá-los e os injetando na SEC, observa-se que essas populações de dímeros são estáveis. Analisando os cromatogramas (**Fig. 26**) pode ser observado que: nessa amostra a 40 °C e 50 °C já se observa uma pequena população de monômeros (79 mL) e uma possível pequena população de tetrâmeros (58 mL), a 60 °C a população monomérica aumenta discretamente. A 70 °C observou-se um aumento razoável na população de monômeros e o desaparecimento da população de tetrâmeros. E após o aquecimento a 80 °C, foi observado um grande aumento na população monomérica acompanhado de uma redução na população de dímeros, e também do surgimento de uma nova população que eluiu em torno de 49 mL, que supomos serem agregados. Esse resultado difere do observado para NOG1, que apresentou aproximadamente a mesma quantidade de monômeros quando aquecida a 70 °C e a 80 °C.

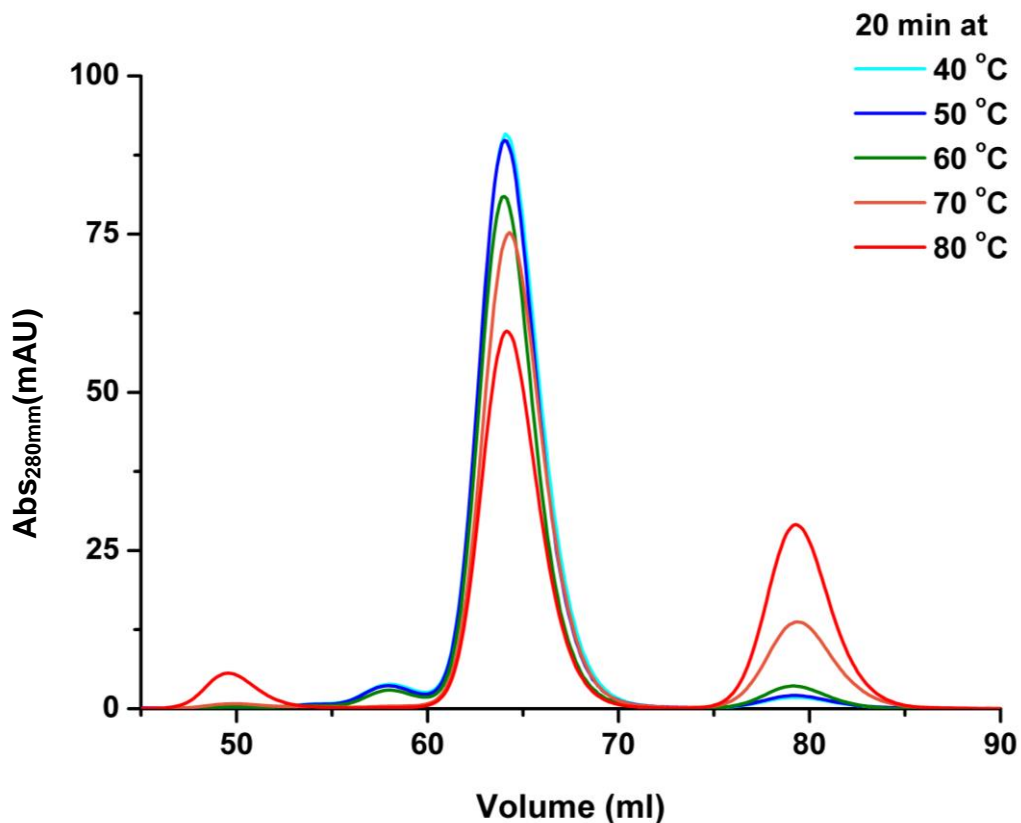


Figura 26. Sobreposição de cromatogramas de exclusão molecular em uma coluna Superdex 75 (10/300) de amostras contendo dímeros de CAF1. As amostras injetadas foram previamente incubadas por 20 min nas temperaturas indicadas. O pico referente a dímeros ocorre em torno de 64 mL e o pico referente a monômeros está em torno de 79 mL. O pico em torno de 49 mL é atribuído a agregados de alto peso molecular. A figura ilustra a influência da temperatura na formação de monômeros a partir de dímeros de CAF1.

Com estes dados, resolvemos aquecer monômeros de NOG1 e submetê-los à cromatografia de exclusão molecular. Aquecendo o uma alíquota contendo 0,62 mg/mL de monômeros de NOG1 nas mesmas temperaturas, por 20 min em banho-maria, e em seguida injetando na SEC, podemos observar que os monômeros são mais instáveis, e nas temperaturas acima de 40 °C parte da população monomérica é convertida em população dimérica. Analisando o cromatograma (**Fig. 27**), observamos que a população de monômeros na quando aquecida a 40 °C não se converte em dímeros, a 50 °C uma pequena população se converte em dímeros, a 60 °C grande parte é convertida em dímeros, e a 70 °C a maioria da proteína se encontra na forma dimérica.

Resultados semelhantes foram obtidos para monômeros de CAF1 na concentração de 0,70 mg/mL (**Fig. 28**).

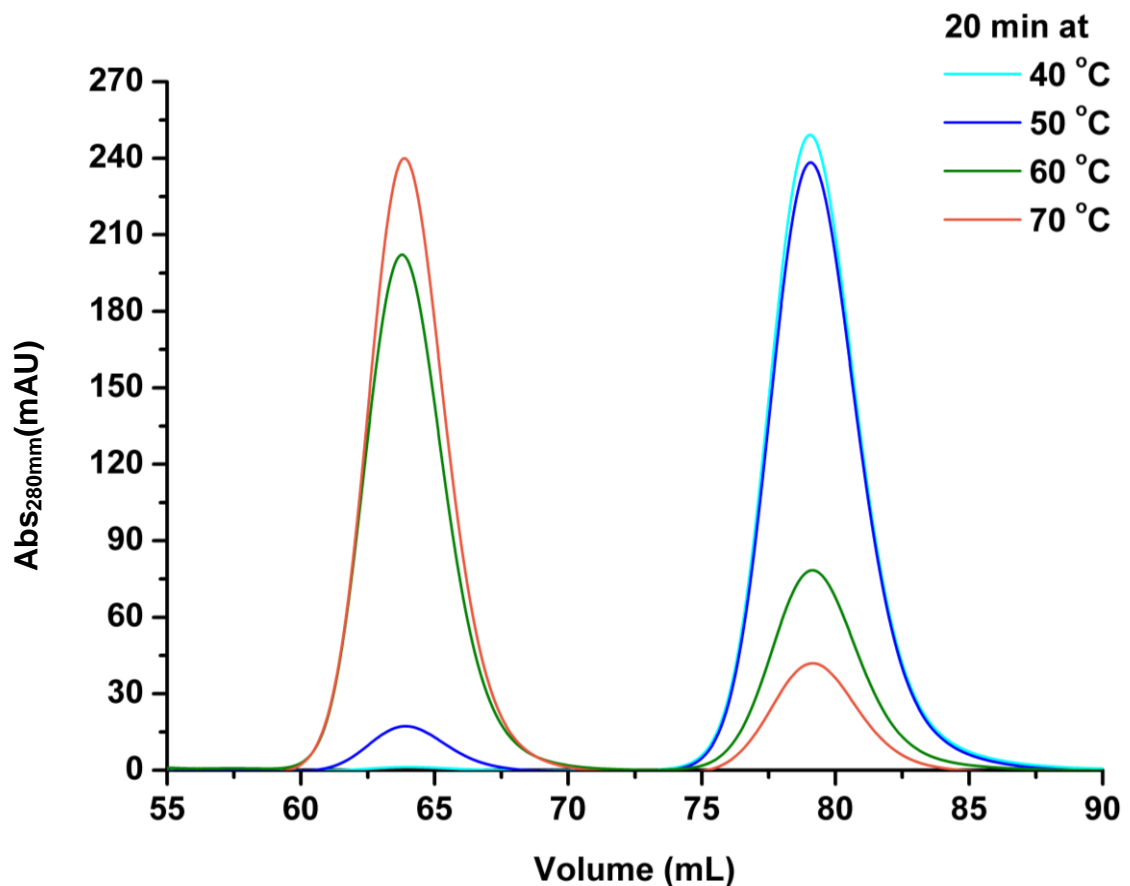


Figura 27. Sobreposição de cromatogramas de SEC ilustrando a influência da temperatura na formação de dímeros a partir de monômeros de NOG1. O pico referente a dímeros ocorre em torno de 64 mL e o pico referente a monômeros está em torno de 79 mL.

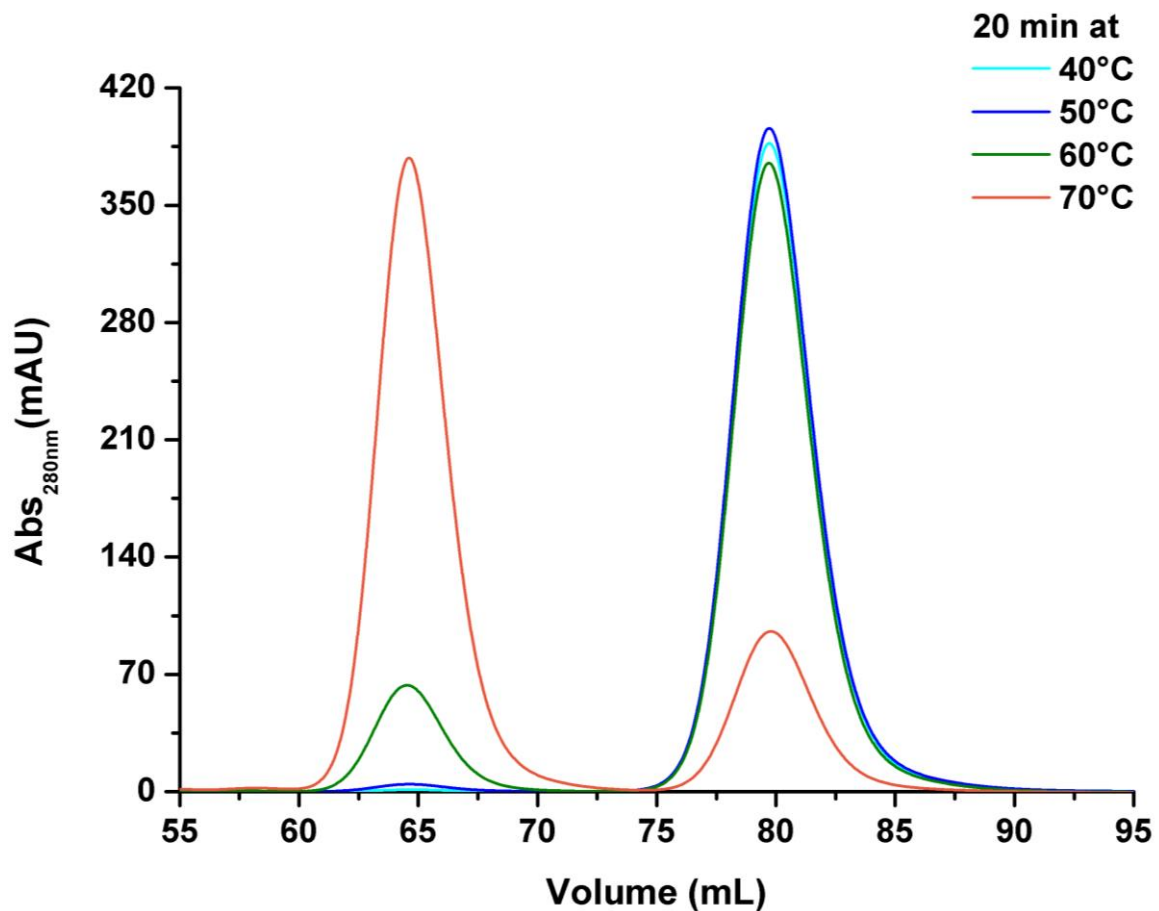


Figura 28. Sobreposição de cromatogramas de SEC ilustrando a influência da temperatura na formação de dímeros a partir de monômeros de CAF1. O pico referente a dímeros ocorre em torno de 65 mL e o pico referente a monômeros está em torno de 80 mL.

Na literatura Freitas e colaboradores (2015), demonstraram que fitocistatinas como as TcCYS3 e TcCYS4 (*Theobroma cacao*) podem ser estáveis na temperatura de 90 °C, além de manterem em mais de sua atividade inibitória na protease papaína, em quanto Moura (2019) em experimentos de ultracentrifugação analítica para observar a formação de oligômeros, foi constatado que a mudança de temperatura alterava as conformações das proteínas estudadas. Nossos resultados corroboram a hipótese de que o dímero é mais estável, possivelmente devido a contribuição entálpica das ligações de hidrogênio intermoleculares na interface aberta, que apenas existe em dímeros. Convém lembrar que esses resíduos representam o motivo QVVSG, que no monômero formam o primeiro loop inibitório.

Outro parâmetro que pode influenciar a predominância de determinada espécie oligomérica em solução é a concentração total de proteína. É razoável supor que em altas concentrações existe uma maior chance de duas moléculas de proteína interagirem, ao passo que em baixíssimas concentrações de proteína dois monômeros possuem uma menor chance de interação, e caso um dímero se converta em monômeros em baixa concentração, a chance de esses monômeros interagirem com outros monômeros será menor, devido a baixa concentração. Para testar essa hipótese foi planejado um experimento utilizando dímeros de NOG1 a uma concentração de 0,40 mg/mL. Essa amostra foi diluída por diluição seriada, e as alíquotas foram aquecidas a 70 °C por 20 min, e injetadas em uma coluna Superdex 75 (10/300) (**Fig. 29**).

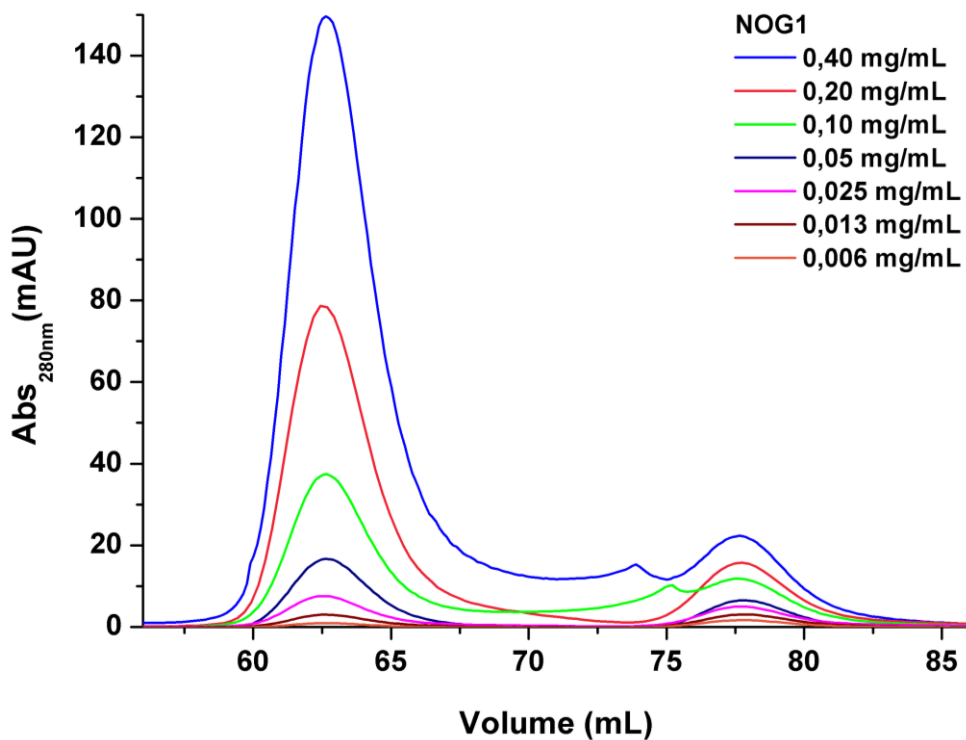


Figura 29. Sobreposição de cromatogramas de SEC ilustrando o equilíbrio de conversão de dímeros/monômeros de NOG1, sob variações de concentrações da proteína. Cada curva corresponde a uma diluição da proteína NOG1 que foi aquecida a 70 °C por 20 min.

Foi observado que quanto maior a concentração, maior a proporção da população de dímeros remanescentes. Para facilitar a visualização dessa observação, a **Figura 30** apresenta a porcentagem do sinal de absorvância de cada curva correspondente ao pico de monômeros e ao de dímeros. Nesse gráfico, cada

conjunto de dois pontos de mesma abscissa (um preto e um vermelho) representa uma curva da **Fig. 29**. A soma dos valores desses dois pontos de mesma abscissa sempre será 100 %. Podemos observar que quanto maior a diluição da amostra de NOG1 injetada, maior a proporção de monômeros na composição do sinal de absorvância no procedimento de SEC.

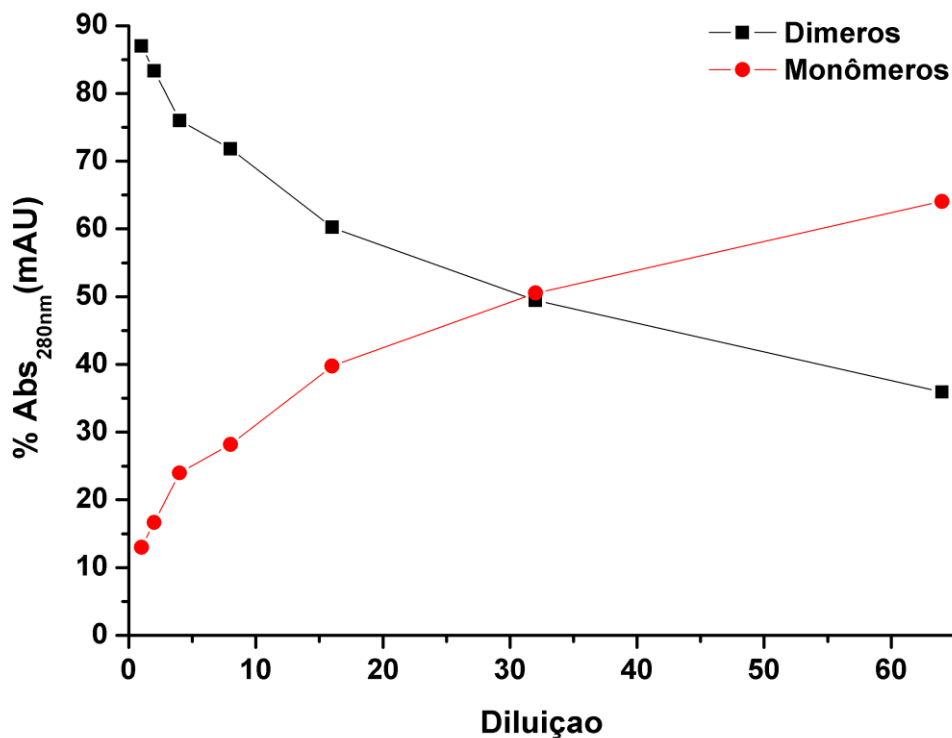


Figura 30. Porcentagem do sinal de absorvância correspondente ao pico de dímeros e ao pico de monômeros. A curva preta representa os valores de porcentagem de absorvância associados aos dímeros e a curva vermelha representa os valores para monômeros. Cada conjunto de dois pontos de mesma abscissa (um preto e um vermelho) representa uma curva da **Fig. 29**.

Nossos resultados demonstram que a interconversão entre monômeros e dímeros é dependente da concentração. Concentrações elevadas favorecem a formação de dímeros. Os ensaios de oligomerização apresentados aqui abordam de forma quantitativa aspectos da dinâmica de interconversão entre monômeros e dímeros por troca de domínio, evento relacionado à dinâmica e estabilidade dessas proteínas.

5.4 Cristalização

Foram realizados vários ensaios de cristalização com as proteínas NOG1 e CAF1 puras, e também dessas proteínas em misturas equimolares com papaína adquirida da empresa Sigma. As proteínas foram utilizadas em diferentes concentrações, variando de 10 mg/mL a mais de 90 mg/mL, essas concentrações elevadas foram possíveis devido a alta solubilidade dessas proteínas, o que pode ser um fator que dificulte a cristalização. O robô de mosquito[®] HTS foi utilizado nos ensaios robotizados iniciais em placas de 96 poços em conjunto com quatro kits comerciais: Crystal Screen (Hampton Research), Morpheus (Molecular Dimensions), JCSGplus (Molecular Dimensions) e PACT Premier (Molecular Dimensions).

A maioria das gotas se manteve transparentes, sem presença de precipitado, indicando que as proteínas estudadas apresentam alta solubilidade mesmo em altas concentrações e na presença de agentes precipitantes. As condições que se mostraram mais promissoras para NOG1 foram: 0,2 M cloreto de magnésio hexahidratado, 0,1 M HEPES pH 7,5, 30% v/v PEG 400; 0,1 M fosfato–citrato pH 4,2, 40% v/v PEG 300; 0,2 M sulfato de lítio, 0,1 M tris pH 8,5, 40% v/v PEG 400.

Foram obtidos cristais nos refinamentos manuais de NOG1 com 0,2 M cloreto de magnésio hexahidratado, 30% v/v (PEG 300 ou PEG 400) e 0,1 M HEPES variando o pH de 7,3 a 8. Porém, estes cristais não difrataram bem. Já os cristais do refinamento manual de fosfato citrato (0,2 M de Na_2HPO_4 , 0,1 M de ácido cítrico, pH 4,4), difrataram a melhores resoluções, chegando a 3,1 Å de resolução (**Fig. 31**).

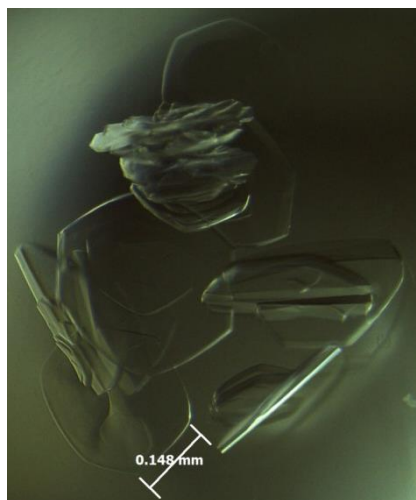


Figura 31. Cristais de NOG1 obtidos em 0,1 M fosfato–citrato pH 4,2, 40% v/v PEG 300.

Foram realizados refinamentos manuais para as seguintes condições de CAF1: 0,2 M brometo de sódio; 20% PEG 3350, Bis-tris propane pH 6,6, e 0,6 M sulfato de amônio, 25% PEG 3350 e 2% de etileno glicol. Contudo, os cristais apresentam forma de agulhas e baixa capacidade de difração.

Cristais de NOG1+Papaína foram obtidos na triagem robotizada em placas de 96 poços, e duas condições foram refinadas manualmente, 0,1 M imidazol, 1 M acetato de sódio pH 6,5, e 0,1 M bicina pH 9, 10% PEG 6000. Um cristal do complexo NOG1+Papaína obtido em 0,1 M imidazol e 1 M acetato de sódio pH 6,5 difratou até aproximadamente 2,2 Å, porém o conjunto de dados apresenta uma série de limitações, que incluem anisotropia, grupo espacial P1 e tamanho da unidade assimétrica muito grande. Não foi possível obter uma solução por substituição molecular apesar das diversas abordagens tentadas.

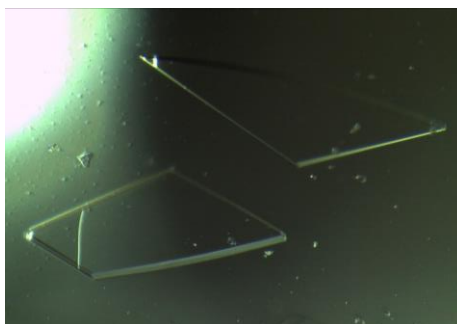


Figura 32. Imagem de cristais do complexo NOG1+Papaína. Os cristais foram obtidos na condição com 0,1 M imidazol, 1 M acetato de sódio pH 6,5. .

5.5 Coleta de Dados de Difração de Raios-X e Resolução da Estrutura

Os experimentos de difração de raios-X foram executados no Laboratório de Biofísica Molecular da Universidade de Brasília. Os cristais obtidos nos ensaios automáticos e manuais de cristalização foram coletados dos poços com o auxílio de loops de tamanho apropriado e posicionados no difratômetro D8-VENTURE (BRUKER) com detector PHOTON II, onde foram crioresfriados sob um feixe de nitrogênio a 100 K. Alguns cristais foram previamente transferidos para uma solução crioprotetora contendo a solução mãe do poço e diferentes concentrações de etilenoglicol ou glicerol.

Muitos cristais foram testados, contudo devido ao atraso de duas semanas no fornecimento de nitrogênio para o início dos experimentos de difração, os cristais de NOG1 já se encontravam relativamente deteriorados quando foram testados. De toda forma, foram coletados 9 conjuntos de dados, e todas as coletas utilizaram o comprimento de onda fixo de 1,54184 Å. Os principais parâmetros variados nessas coletas foram o número de imagens (de 360 a 1440), a oscilação (de 0,25° a 1°) e a distância do detector (de 80 cm a 150 cm) e o tempo de exposição. Contudo, vários desses conjuntos de dados apresentaram diversas limitações, sendo a principal baixa resolução.

Após o processamento dos dados, foi identificado o melhor conjunto de dados (Nogx3), contendo 1440 imagens, com tempo de exposição de 20 segundos por imagem, coletadas com oscilação de 0,25° e com o detector a uma distância de 80 cm. Esse cristal foi obtido na condição contendo 0,2 M de Na₂HPO₄ e 0,1 M de ácido cítrico, pH 4,4, 40 % v/v PEG 300. O programa XDS foi utilizado para indexar, integrar e escalonar as reflexões desse conjunto de dados. O grupo espacial foi determinado como C 1 2 1, e as dimensões da célula unitária são a = 106,77, b = 79,35 e c = 63,91, com os ângulos $\alpha = 90,00$, $\beta = 108,11$ e $\gamma = 90,00$. Utilizando como parâmetros de definição da resolução do conjunto de dados na última faixa de resolução completudeza > 96 %, $I/\sigma > 1$ e $CC_{1/2} > 60$ (Karplus e Diederichs, 2012), a resolução foi definida como 3,1 Å. O programa PHENIX XTRIAGE e o servidor STARANISO indicam que o conjunto de dados apresenta moderada anisotropia, o que pode ser visto de forma simplificada como uma diferença na capacidade de difração do cristal em diferentes direções.

O procedimento de substituição molecular realizado no programa PHASER usou como *template* na busca um modelo de monômero da cistatina do lúpulo (*Humulus lupulus*), produzido a partir do PDB 6VLP. O coeficiente de Matthews indica a presença de 4 cópias da proteína NOG1 na unidade assimétrica e um conteúdo de solvente de aproximadamente 53 %. Utilizando as fases desse modelo, o programa PHASER foi capaz de identificar uma solução com densidade eletrônica plausível. Essa solução foi posteriormente refinada utilizando os programas PHENIX e COOT, e a redução consecutiva nos valores de R_{work} e R_{free} indicam que a solução é correta.

No início do processo de refinamento foi observado que as quatro cópias da NOG1 estão presentes na unidade assimétrica na forma de dois dímeros *domain swap*. O modelo foi adequado e o refinamento teve continuidade, sendo que a cada ciclo de refinamento os valores de R_{work} e R_{free} foram monitorados, e o servidor Molprobitry foi constantemente consultado para avaliar a estereoquímica do modelo resultante.

As estatísticas de coleta de dados e refinamento da estrutura de NOG1 estão apresentadas na **Tabela 3**.

Devido a uma série de limitações, das quais a principal é a resolução limitada a 3,1 Å e outra relevante é a presença de anisotropia nos dados identificada pelo programa PHENIX_XTRIAGE, que indica que o padrão de difração apresenta resoluções diferentes em diferentes direções. Se possível, a interpretação do modelo cristalográfico deve ser feita sempre em conjunto com a densidade eletrônica, e tendo em vista as limitações apresentadas acima. A densidade para diversas cadeias laterais é ambígua, contudo a densidade eletrônica deixa explícito que são observados na unidade assimétrica dois dímeros por troca de domínios.

Tabela 3 – Estatísticas de coleta e refinamento de dados para o conjunto de dados NOG1. As estatísticas para a camada de maior resolução são mostradas entre parênteses.

Grupo espacial	C 1 2 1
Faixa de resolução (Å)	30,59 – 3,104 (3,215 – 3,104)
Dimensões da célula unitária a; b e c (Å)	106,767 79,346 63,914 90 108,114 90
Total de reflexões	67904 (6537)
Reflexões únicas	9204 (904)
Multiplicidade	7,4 (7,2)
Completeza (%)	98,51 (96,99)
Média I/σ (I)	7,07 (1,20)
CC_{1/2}	0,986 (0,721)
Reflexões utilizadas no refinamento	9144 (901)
Reflexões utilizadas para o R_{free}	459 (46)
R_{work}	0,2551
R_{free}	0,3029
Número de átomos da proteína	2660
Número de átomos do solvente	ainda sem H ₂ O
Resíduos de proteína	327
RMS (ligações)	0,011
RMS (ângulos)	1,36
Ramachandran favorecidos (%)	83,07
Ramachandran permitidos (%)	14,11
Ramachandran outliers (%)	2,82
Outliers de rotâmeros (%)	0,35
Clash score	20,41
Média B-factor	48,89

5.6 Estrutura Cristalográfica da NOG1

Assim como esperado e corroborando com os resultados da cromatografia de exclusão molecular, o modelo refinado de NOG1 apresenta dois dímeros por troca de domínios na unidade assimétrica. O RMSD entre os dois dímeros *domain swap* é de 0,586 (**Fig. 33**). Até o momento, o modelo refinado possui densidade eletrônica suficiente para modelar todas as 4 cadeias do resíduo 17 ao resíduo 98.

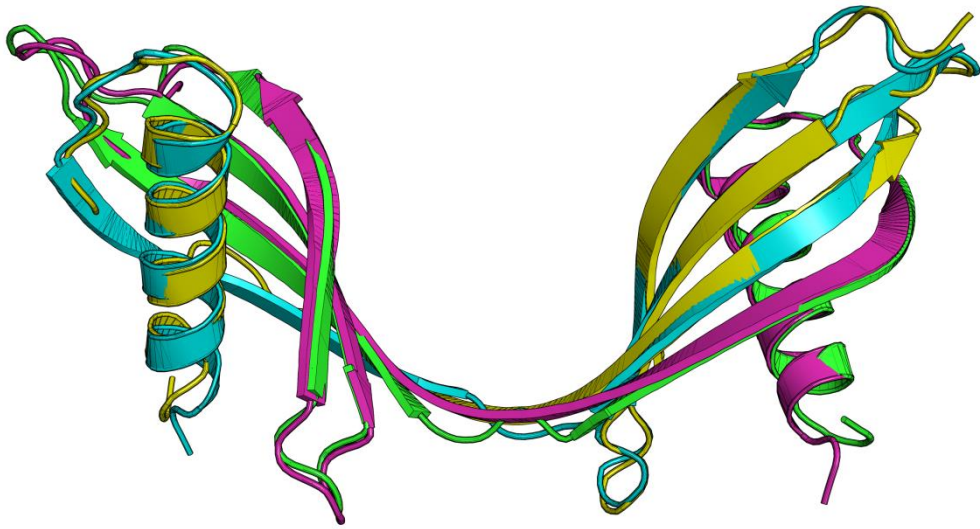


Figura 33. Sobreposição dos dois dímeros por troca de domínios (cadeias AB e CD) contidos na unidade assimétrica da estrutura de NOG1. O valor de RMSD da sobreposição é de 0,586. Cadeia A está apresentada em verde, cadeia B em ciano, cadeia C em magenta e cadeia D em amarelo. A cadeia A está sobreposta a cadeia C e a cadeia B está sobreposta a cadeia D.

Na unidade assimétrica os dois dímeros (AB e CD) estão relacionados por simetria não cristalográfica (**Fig. 34**), e a cadeia A faz com a cadeia D interações correspondentes às interações que a cadeia C faz com a cadeia B. Um exemplo dessas interações é o posicionamento do nitrogênio da cadeia lateral de Lys79 das cadeias B e D próximo ao dipolo negativo das hélices das cadeias C e A, respectivamente, fazendo uma ligação de hidrogênio com o oxigênio da cadeia principal de Lys33 (**Fig. 35**).

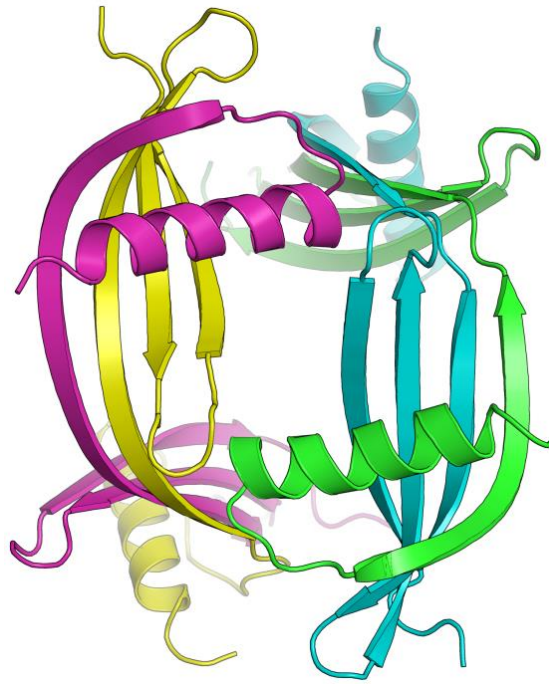


Figura 34. Conteúdo da unidade assimétrica da estrutura NOG1. Cada cadeia polipeptídica está representada em uma cor, e um dímero por troca de domínios colorido em verde e ciano se relaciona por simetria não cristalográfica com o outro dímero por troca de domínios colorido em magenta e amarelo.

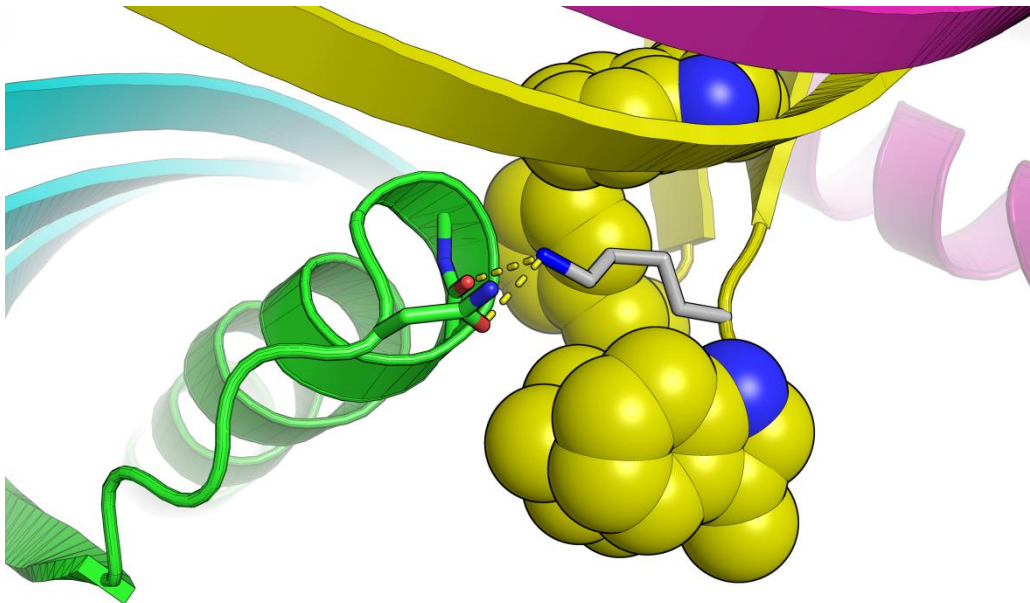


Figura 35. Interação entre os dímeros na unidade assimétrica. O nitrogênio da cadeia lateral da Lys79 da cadeia D se posiciona próximo ao dipolo negativo da hélice da cadeia A, e faz ligações de hidrogênio com o oxigênio da cadeia principal de Lys33 e com o oxigênio da cadeia lateral de Asn36. As cadeias laterais dos resíduos Trp81, Leu82, Phe84 e Trp77 da cadeia D, representadas como esferas, criam um ambiente hidrofóbico que favorece o posicionamento da cadeia lateral da Lys79. Cadeia A está apresentada em verde, cadeia B em ciano, cadeia C em magenta e cadeia D em amarelo.

O enovelamento das cistatinas apresenta uma folha β antiparalela curvada em torno de uma hélice. No caso de dímeros por troca de domínios, temos dois desses lobos conectados por duas extensas fitas antiparalelas, cada uma proveniente de uma cadeia polipeptídica. Dependendo de como é medido, o ângulo entre as hélices na estrutura de NOG1 varia de 73° a 84° para as cadeias AB e de 76° a 86° para as cadeias CD (**Figura 36**), valores relativamente grandes. Para fins de comparação, os valores de ângulos entre as hélices da cistatina do lúpulo variam entre 46° e 62° , e os da cistatina da *Cannabis sativa* variam de 68° a 77° .

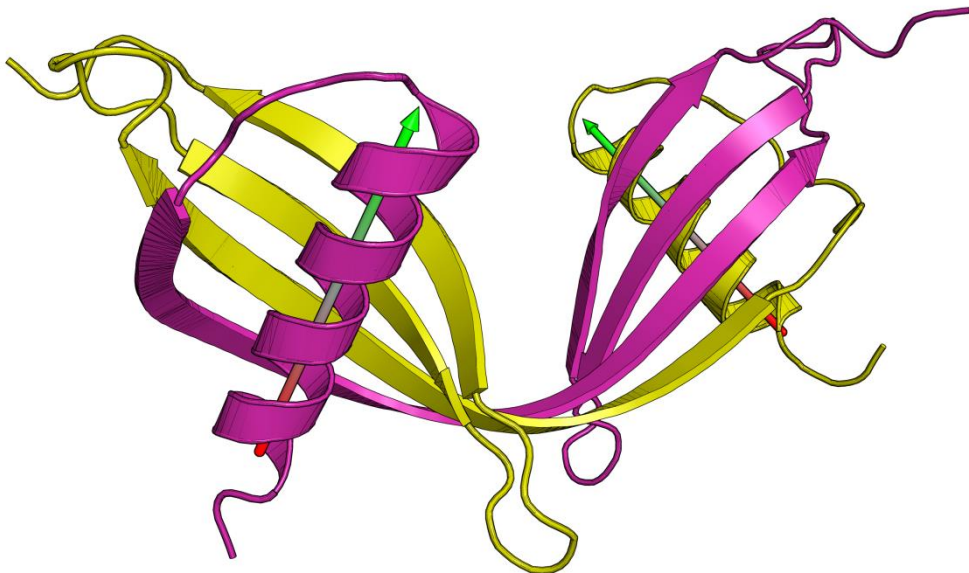


Figura 36. Ângulos entre as hélices em NOG1. Devido à troca de domínios, cada lobo apresenta componentes das cadeias C (magenta) e D (amarelo). As setas no centro das hélices foram utilizadas para definir os ângulos.

Até o momento da escrita dessa dissertação o refinamento está bastante avançado, contudo ainda não está concluído. Os valores de R_{work} e R_{free} são 0,256 e 0,302, respectivamente. Esses valores são aceitáveis para uma estrutura de resolução 3,1 Å (Kleywegt e Jones, 2002) e moderada anisotropia. O servidor MolProbity (<http://molprobity.biochem.duke.edu/>) atribui ao modelo atual o percentil 96 no *Clashscore* e também percentil 96 no *MolProbity score*, uma análise que indica que a estereoquímica desse modelo se encontra entre as melhores estruturas de resolução comparável. Apesar desses indicadores, a densidade eletrônica apresenta limitações, como pode ser observado na **Figura 37**. Mesmo tendo em

vistas as limitações impostas pela resolução, a estrutura da NOG1 é muito importante no contexto do presente trabalho, principalmente porque são observados dímeros na unidade assimétrica. Os ensaios de cristalização utilizaram amostras de dímeros que foram mantidos sempre a baixas temperaturas, um fator crucial para evitar a formação de monômeros, que foi identificado nos ensaios em solução. Amostras de monômeros não apresentaram formação de cristais na mesma condição de cristalização, indicando que a interconversão entre esses estados oligoméricos não ocorre nessas condições.

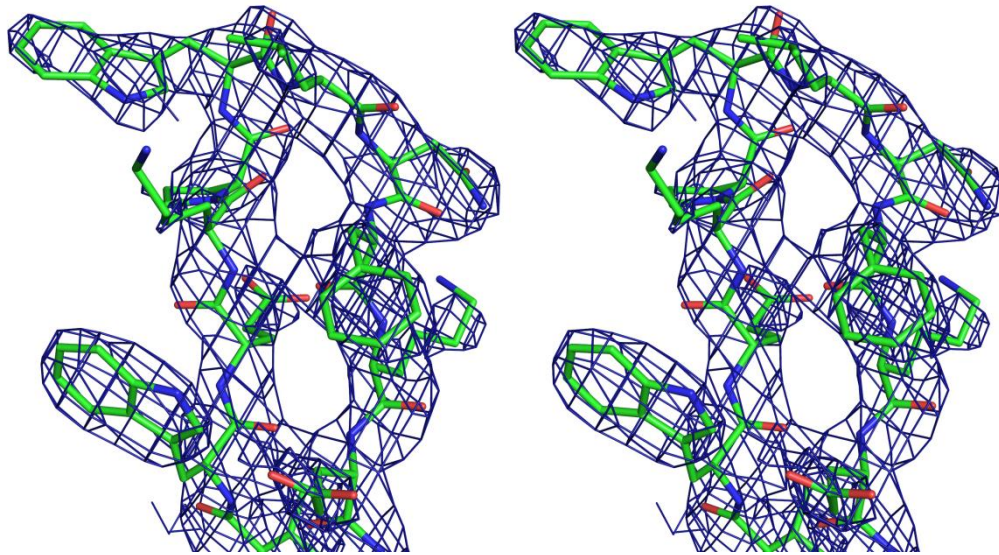


Figura 37. Imagem stéreo do mapa de densidade eletrônica $2F_o-F_c$ contornado em 2σ para os resíduos 77 a 86 (sequência: IWEKPWLNFKKE), que incluem o motivo PW do segundo loop inibidor.

5.7 Empacotamento Cristalino

O cristal monoclinico de grupo espacial $C 1 2 1$ contém quatro cópias de NOG1 em sua unidade assimétrica, organizadas na forma de dois dímeros por troca de domínios. O conteúdo de solvente do cristal, estimado pelo coeficiente de Mathews, é de 53%. Canais de solvente interconectados em todas as direções são observados, como ilustrado na **Figura 38**.

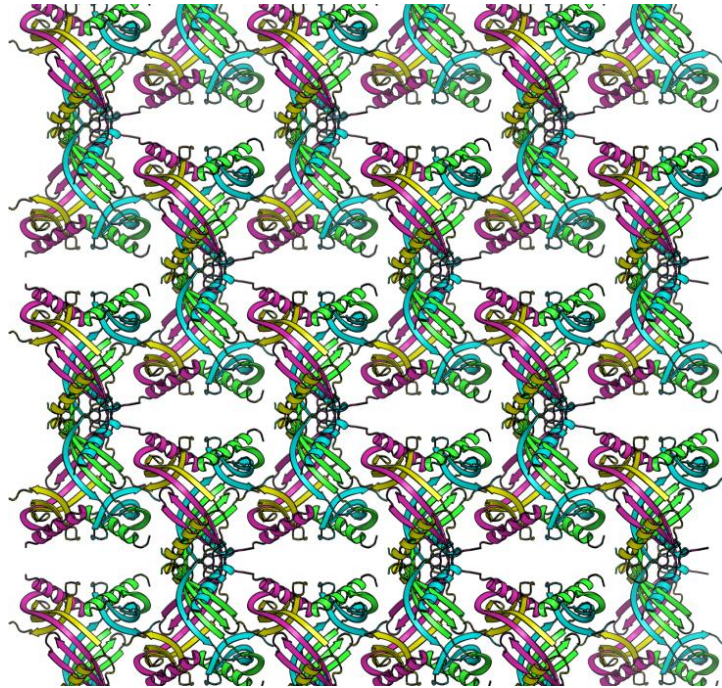


Figura 38. Imagem ortoscópica ilustrando o empacotamento cristalino da estrutura da NOG1. Com um conteúdo de solvente de 53%, podem ser observados grandes canais de solvente no arranjo cristalino. Cadeia A está apresentada em verde, cadeia B em ciano, cadeia C em magenta e cadeia D em amarelo.

5.8 Fitocistatinas e Biotecnologia

A atividade inibitória das fitocistatinas regula processos da fisiologia da planta, e essas proteínas também são capazes de inibir proteases de insetos herbívoros e de patógenos. Assim, as fitocistatinas atuam como versáteis reguladoras de cisteíno proteases endógenas e exógenas. O presente trabalho elucida algumas das condições onde a regulação dessa atividade inibitória pela formação de dímeros por troca de domínios ocorre.

Uma aplicação biotecnológica potencial pode ser o desenvolvimento de plantas com fitocistatinas modificadas ou de outros organismos visando, por exemplo, inibir cisteíno proteases presentes no trato digestivos de insetos herbívoros (Arai *et al.*, 2002; Goulet *et al.*, 2008; Martinez *et al.*, 2016), dificultando a digestão, que por sua vez, afeta o crescimento e reprodução dos insetos (Dunse *et al.*, 2010; Parde *et al.*, 2012; War *et al.*, 2012; Jadhav *et al.*, 2016). Alguns estudos já demonstraram que esse mecanismo de defesa através das PhyCys é efetivo contra insetos das ordens coleóptera (Zhu–Salzman *et al.*, 2003; Ahn *et al.*, 2007), lepidóptera (Parde *et al.*, 2012; Jadhav *et al.*, 2016) e hemiptera (Azzouz *et al.*, 2005). As principais pragas da noqueira e do café estão apresentadas na **Tabela 4**.

Diversos insetos prejudicam o desenvolvimento de plantas como o café e a noqueira. Algumas das pragas do café são espécies de insetos dentro das ordens coleoptera, lepidóptera e hemiptera (Mesquita *et al.*, 2016). Outros insetos pertencentes a essas mesmas ordens também foram identificados como pragas para diversas espécies de noqueiras (Khan *et al.*, 2016). Sendo assim, as PhyCys presentes em ambas as plantas representam possíveis mecanismos de proteção contra herbivoria dessas ordens de insetos e parte do nosso estudo pode auxiliar no desenvolvimento de novos bioinseticidas.

Tabela 4 – Principais pragas das plantas café e noqueira. Nome científico em itálico e entre parênteses ordem ao qual pertence.

Nogueira	Café
<i>Chromaphis juglandicola</i> (hemiptera)	<i>Leucoptera coffeella</i> (lepidóptera)
<i>Callipteras Juglandis</i> (hemiptera)	<i>Hypothenemus hampei</i> (coleoptera)
<i>Conotrachelus juglandis</i> (coleoptera)	<i>Quesada gigas</i> (hemiptera)
<i>Quadraspidotus juglansregiae</i> (hemiptera)	<i>Dysmicoccus texensis</i> (hemiptera)
<i>Lymantria obfusca</i> (lepidoptera)	<i>Eacles imperialis magnifica</i> (lepidoptera)
<i>Adoretus simplex</i> (coleoptera)	
<i>Scolytus nitidus</i> (coleoptera)	
<i>Rhagoletis completa</i> (Diptera)	

Também é possível garantir uma maior produção através do melhoramento genético de plantas e uma das formas de realizar o melhoramento genético é através da transgenia. Com a aplicação biotecnológica na produção de plantas transgênicas é possível alterar diversas características, principalmente aquelas relacionadas à tolerância de estresses abióticos, através da regulação da expressão das PhyCys.

Tan e colaboradores (2015) utilizaram plantas transgênicas expressando PhyCys exógenas para demonstrar sua relação quanto à resistência à seca, proteção contra *stress* oxidativo devido à desidratação (Tan *et al.*, 2015) e tolerância a elevadas taxas de salinidade no solo (Tan *et al.*, 2016). PhyCys também estão envolvidas no desenvolvimento das plântulas e na germinação das sementes (Münt, 1996; Toyooka *et al.*, 2000; Hwang *et al.*, 2009; Szewińska *et al.*, 2016), inclusive já foi verificado que a expressão heteróloga da fitocistatina BrCYS1 (*Brassica rapa* fitocistatina 1) é capaz de retardar ou acelerar a germinação de sementes (Hong *et al.*, 2007), fatores que podem contribuir no cultivo, garantindo que a germinação ocorra durante períodos mais favoráveis para o desenvolvimento da planta.

Os resultados do presente estudo contribuem para o melhor entendimento das propriedades das fitocistatinas e podem auxiliar no desenvolvimento de novas aplicações biotecnológicas para as PhyCys da noqueira, do café ou de outras plantas.

6 CONCLUSÃO

No presente trabalho apresentamos uma metodologia eficiente para a produção das fitocistatinas NOG1 e CAF1, provenientes da noqueira e do cafeeiro, respectivamente. Utilizando protocolos robustos de expressão heteróloga em meio auto-indutor, aliados a etapas de purificação por afinidade a níquel e purificação por cromatografia de exclusão molecular, foi possível obter ambas fitocistatinas com grau de pureza alto o suficiente para proceder com ensaios de cristalização. Os cristais obtidos foram direcionados para experimentos de difração *in-house* no difratômetro D8-Venture, diversos conjuntos de dados foram coletados e a estrutura cristalográfica da proteína NOG1 foi resolvida por substituição molecular e refinada a valores de R_{work} e R_{free} compatíveis com a resolução de 3,1 Å. A estrutura apresenta dois dímeros por troca de domínios na unidade assimétrica.

Em paralelo, foram identificadas condições nas quais é possível armazenar por longos períodos monômeros e dímeros de NOG1 e CAF1 sem que haja interconversão significativa entre esses estados oligoméricos. A partir daí, foi possível desenvolver ensaios capazes de avaliar algumas das condições que propiciam a conversão de monômeros em dímeros por troca de domínios, e vice-versa. Foi observado que concentrações maiores favorecem a formação de dímeros, principalmente se as amostras forem expostas a temperaturas acima de 50 °C por um intervalo de tempo de 20 min ou mais.

A análise estrutural em conjunto com os dados de oligomerização permitiu um entendimento inédito do processo de dimerização dessas proteínas, um conhecimento que pode ser extrapolado para outras fitocistatinas.

Em conjunto, as metodologias e resultados apresentados consolidam a tecnologia de produção e isolamento de monômeros e dímeros de fitocistatinas no Laboratório de Biofísica Molecular da UnB. Uma das perspectivas é identificar os determinantes estruturais responsáveis pela predominância de um estado oligomérico em solução, como monômeros, e aplicar essas informações no desenvolvimento de ferramentas úteis do ponto de vista biotecnológico.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ABE, K.; EMORI, Y.; KONDO, H.; SUZUKI, K. *et al.* Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice seeds. **Journal of Biological Chemistry**, 262, n. 35, p. 16793–16797, 1987.

ABE, K.; KONDO, H.; ARAI, S. Purification and characterization of a rice cysteine proteinase inhibitor. **Agricultural and Biological Chemistry**, 51, n. 10, p. 2763–2768, 1987.

AHN, J.-E. *et al.* Cowpea bruchid *Callosobruchus maculatus* counteracts dietary protease inhibitors by modulating propeptides of major digestive enzymes. **Insect Molecular Biology**, v. 16, n. 3, p. 295-304, 2007.

ALBERTS, B.; BRAY, D.; HOPKIN, K.; JOHNSON, A. D. *et al.* **Essential cell biology**. Garland Science, 2013. 1317806271.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W. *et al.* Basic local alignment search tool. *Journal of molecular biology*, 215, n. 3, p. 403-410, 1990.

ARAI, S.; MATSUMOTO, I.; EMORI, Y.; ABE, K. Plant seed cystatins and their target enzymes of endogenous and exogenous origin. **Journal of Agricultural and Food Chemistry**, 50, n. 22, p. 6612–6617, 2002.

AZZOUZ, Hichem *et al.* Effects of plant protease inhibitors, oryzacystatin I and soybean Bowman–Birk inhibitor, on the aphid *Macrosiphum euphorbiae* (Homoptera, Aphididae) and its parasitoid *Aphelinus abdominalis* (Hymenoptera, Aphelinidae). **Journal of insect physiology**, v. 51, n. 1, p. 75-86, 2005.

BACKES, P.; IRGANG, B. E. **Árvores do Sul: guia de identificação & interesse ecológico**. Paisagem do Sul, 2009.

BARRETT, A. J. [1] Classification of peptidases. *In: Methods in enzymology*: Elsevier, 1994. v. 244, p. 1–15.

BENCHABANE, M.; SCHLÜTER, U.; VORSTER, J.; GOULET, M.–C. *et al.* Plant cystatins. **Biochimie**, 92, n. 11, p. 1657–1666, 2010.

BENNETT, M. J.; EISENBERG, D. The evolving role of 3D domain swapping in proteins. **Structure**, 12, n. 8, p. 1339-1341, 2004.

BENNETT, M. J.; SCHLUNEGGER, M. P.; EISENBERG, D. 3D domain swapping: a mechanism for oligomer assembly. **Protein Science**, 4, n. 12, p. 2455-2468, 1995.

BENNETT, M.; CHOE, S.; EISENBERG, D. Domain swapping: entangling alliances between proteins. **Proceedings of the National Academy of Sciences**, 91, n. 8, p. 3127-3131, 1994.

BERROW, N. S.; BÜSSOW, K.; COUTARD, B.; DIPROSE, J. *et al.* Recombinant protein expression and solubility screening in *Escherichia coli*: a comparative study. **Acta Crystallographica Section D: Biological Crystallography**, 62, n. 10, p. 1218-1226, 2006.

BERRY, E. W. Notes on the Geological History of the Walnuts and Hickories. **The Plant World**, 15, n. 10, p. 225–240, 1912.

BEVAN, L.; STRETTON, C. Effects of Secondary Interactions in Size Exclusion Chromatography.

BRIDSON, D.; VERDCOURT, B. *Coffea*. **Flora of Tropical East Africa**, Rubiaceae, 2, p. 703-723, 1988.

BRZIN, J.; KOPITAR, M.; LOČNIKAR, P.; TURK, V. An endogenous inhibitor of cysteine and serine proteinases from spleen. **FEBS letters**, 138, n. 2, p. 193–197, 1982.

BUSSO, D.; STIERLÉ, M.; THIERRY, J.-C.; MORAS, D. A comparison of inoculation methods to simplify recombinant protein expression screening in *Escherichia coli*. **Biotechniques**, 44, n. 1, p. 101–106, 2008.

CALLENDER, R.; DYER, R. B. Advances in time-resolved approaches to characterize the dynamical nature of enzymatic catalysis. **Chemical reviews**, 106, n. 8, p. 3031–3042, 2006.

CALLENDER, R.; DYER, R. B. The dynamical nature of enzymatic catalysis. **Accounts of chemical research**, 48, n. 2, p. 407–413, 2015.

CAMPOS, L. H. R. D. Estimativa do progresso genético: Um exemplo aplicado em 18 anos do melhoramento de soja. 2019.

CARLINI, Célia R.; GROSSI-DE-SÁ, Maria Fátima. Plant toxic proteins with insecticidal properties. A review on their potentialities as bioinsecticides. **Toxicon**, v. 40, n. 11, p. 1515–1539, 2002.

CLANCY, S.; BROWN, W. Translation: DNA to mRNA to protein. **Nature Education**, 1, n. 1, p. 101, 2008.

COBB, N. J.; SUREWICZ, W. K. Prion diseases and their biochemical mechanisms. **Biochemistry**, 48, n. 12, p. 2574-2585, 2009.

CONSELHO DOS EXPORTADORES DE CAFÉ DO BRASIL. Relatório de exportações. <http://www.sapc.embrapa.br/arquivos/consorcio/informe_estatistico/CE_CAFE_Relatorio_Mensal_JANEIRO_2020.pdf>. Acesso em 25 de fevereiro de 2020.

CONSELHO DOS EXPORTADORES DE CAFÉ DO BRASIL. Relatório de exportações <<http://www.consorciopesquisacafe.com.br/index.php/imprensa/noticias/959-2019-11-29-17-49-40>>. Acesso em 25 de fevereiro de 2020.

DANDEKAR, A, LESLIE, C and MCGRANAHAN, G, “Juglans regia Walnut,” LITZ, E, R, ed., **Biotechnology of Fruit and Nut Crops** (Cambridge, MA: CABI Pub., 2005): 307

DIOP, N. N.; KIDRIČ, M.; REPELLIN, A.; GAREIL, M. *et al.* A multicystatin is induced by drought-stress in cowpea (*Vigna unguiculata* (L.) Walp.) leaves. **FEBS letters**, 577, n. 3, p. 545-550, 2004.

DOBSON, C. M. Protein folding and misfolding. **Nature**, 426, n. 6968, p. 884–890, 2003.

DUNSE, K. M. *et al.* Coexpression of potato type I and II proteinase inhibitors gives cotton plants protection against insect damage in the field. **Proceedings of the National Academy of Sciences**, v. 107, n. 34, p. 15011-15015, 2010.

DUTT, S.; SINGH, V. K.; MARLA, S. S.; KUMAR, A. In silico analysis of sequential, structural and functional diversity of wheat cystatins and its implication in plant defense. **Genomics, proteomics & bioinformatics**, 8, n. 1, p. 42–56, 2010.

EMBRAPA CAFÉ NOTÍCIAS. Brasil – Exportações dos Cafés do Brasil geram US\$ 438 milhões de receita cambial em janeiro de 2020. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/50174086/exportacoes-dos-cafes-do-brasil-geram-us-438-milhoes-de-receita-cambial-em-janeiro-de-2020>>. Acesso em 25 de fevereiro de 2020.

FAZUOLI, L. C. 1986. Genética e melhoramento do cafeeiro. In: Cultura do Cafeeiro. Factores que afectam a produtividade, **Associação Brasileira para Pesquisa da Potassa e do Fosfato**, São Paulo, pp. 89-113.

FREITAS, Ana Camila Oliveira *et al.* The Activity of Tc CYS4 Modified by Variations in pH and Temperature Can Affect Symptoms of Witches’ Broom Disease of Cocoa, Caused by the Fungus *Moniliophthora perniciosa*. **PLoS one**, v. 10, n. 4, p. e0121519, 2015.

GOULET, M.-C.; DALLAIRE, C.; VAILLANCOURT, L.-P.; KHALF, M. *et al.* Tailoring the specificity of a plant cystatin toward herbivorous insect digestive cysteine proteases by single mutations at positively selected amino acid sites. **Plant physiology**, 146, n. 3, p. 1010–1019, 2008.

GULL, A.; LONE, A. A.; WANI, N. U. I. Biotic and Abiotic Stresses in Plants. *In: Abiotic and Biotic Stress in Plants*: IntechOpen, 2019.

HASHIMOTO, K.; PANCHENKO, A. R. Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states. **Proceedings of the National Academy of Sciences**, 107, n. 47, p. 20352-20357, 2010.

HIBINO, T.; FUKUYAMA, K.; EPSTEIN, W. L. In vitro and in vivo inhibition of rat liver cathepsin L by epidermal proteinase inhibitor. **Biochemical and biophysical research communications**, 93, n. 2, p. 440–447, 1980.

HONG, J. K.; HWANG, J. E.; LIM, C. J.; YANG, K. A. *et al.* Over-expression of Chinese cabbage phytocystatin 1 retards seed germination in Arabidopsis. **Plant science**, 172, n. 3, p. 556–563, 2007.

HOPKINS, W. G. **Introduction to plant physiology**. John Wiley and Sons, 1999. v. Ed. 2). 0471192813.

HWANG, J. E.; HONG, J. K.; JE, J. H.; LEE, K. O. *et al.* Regulation of seed germination and seedling growth by an Arabidopsis phytocystatin isoform, AtCYS6. **Plant cell reports**, 28, n. 11, p. 1623–1632, 2009.

JADHAV, Abhilash R. *et al.* Capsicum annum proteinase inhibitor ingestion negatively impacts the growth of sorghum pest Chilo partellus and promotes differential protease expression. **Biochemistry and biophysics reports**, v. 8, p. 302-309, 2016.

JAENICKE, R.; LILIE, H. Folding and association of oligomeric and multimeric proteins. **Advances in protein chemistry**, 53, p. 329-401, 2000.

JÄRVINEN, M.; RÄSÄNEN, O.; RINNE, A. The low-molecular-weight SH-protease inhibitor in rat skin is epidermal. **Journal of Investigative Dermatology**, 71, n. 2, p. 119–121, 1978.

JÄRVINEN, M.; RINNE, A. Human spleen cysteineproteinase inhibitor: purification, fractionation into isoelectric variants and some properties of the variants. **Biochimica et Biophysica Acta (BBA)–Protein Structure and Molecular Enzymology**, 708, n. 2, p. 210–217, 1982.

JUNIOR, J. E. M.; VALADARES, N. F.; PEREIRA, H. D. M.; DYSZY, F. H. *et al.* Expression in Escherichia coli of cysteine protease inhibitors from cowpea (Vigna unguiculata): The crystal structure of a single-domain cystatin gives insights on its thermal and pH stability. **International journal of biological macromolecules**, 102, p. 29–41, 2017.

KABSCH, W. Xds. **Acta Crystallographica Section D: Biological Crystallography**, 66, n. 2, p. 125–132, 2010.

KHAN, Akhtar Ali *et al.* Insect pests of walnut and their management. **Insect pests management of fruit crops**. Pandey, AK and Mall, P.(Eds). **Biotech, India**, p. 308-355, 2016.

KIDRIĆ, M.; KOS, J.; SABOTIĆ, J. Proteases and their endogenous inhibitors in the plant response to abiotic stress. **Botanica serbica**, 38, n. 1, p. 139-158, 2014.

KIGGUNDU, A.; MUCHWEZI, J.; VAN DER VYVER, C.; VILJOEN, A. *et al.* Deleterious effects of plant cystatins against the banana weevil Cosmopolites sordidus. **Archives of Insect Biochemistry and Physiology: Published in**

Collaboration with the Entomological Society of America, 73, n. 2, p. 87–105, 2010.

KOHEN, A. Role of dynamics in enzyme catalysis: substantial versus semantic controversies. **Accounts of chemical research**, 48, n. 2, p. 466–473, 2015.

KOHLER, Robert. The background to Eduard Buchner's discovery of cell-free fermentation. **Journal of the History of Biology**, p. 35–61, 1971.

KOMINAMI, E.; WAKAMATSU, N.; KATUNUMA, N. Purification and characterization of thiol proteinase inhibitor from rat liver. **Journal of Biological Chemistry**, 257, n. 24, p. 14648–14652, 1982.

KONDO, H.; ABE, K.; EMORI, Y.; ARAI, S. Gene organization of oryzacystatin-II, a new cystatin superfamily member of plant origin, is closely related to that of oryzacystatin-I but different from those of animal cystatins. **FEBS letters**, 278, n. 1, p. 87–90, 1991.

KORDIŠ, D.; TURK, V. Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. **BMC evolutionary biology**, 9, n. 1, p. 266, 2009.

KRIST, S. Coffee Seed Oil. In: *Vegetable Fats and Oils*: Springer, 2020. p. 255-259.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, 227, n. 5259, p. 680, 1970.

LIEBSCHNER, D.; AFONINE, P. V.; BAKER, M. L.; BUNKÓCZI, G. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology*, 75, n. 10, p. 861-877, 2019.

LODISH, H.; BERK, A.; ZIPURSKY, S. L.; MATSUDAIRA, P. *et al.* Molecular cell biology 4th edition. **National Center for Biotechnology Information, Bookshelf**, 2000.

M LIMA, A.; P DOS REIS, S.; RB DE SOUZA, C. Phytocystatins and their potential to control plant diseases caused by fungi. **Protein and peptide letters**, 22, n. 2, p. 104–111, 2015.

MARGIS, R.; REIS, E. M.; VILLERET, V. Structural and phylogenetic relationships among plant and animal cystatins. **Archives of Biochemistry and Biophysics**, 359, n. 1, p. 24–30, 1998.

MARTÍNEZ, M.; CAMBRA, I.; GONZÁLEZ-MELENDI, P.; SANTAMARÍA, M. E. *et al.* C1A cysteine-proteases and their inhibitors in plants. **Physiologia Plantarum**, 145, n. 1, p. 85–94, 2012.

MARTINEZ, M.; DIAZ, I. The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. **BMC Evolutionary Biology**, 8, n. 1, p. 198, 2008.

MARTINEZ, M.; SANTAMARIA, M. E.; DIAZ-MENDOZA, M.; ARNAIZ, A. *et al.* Phytocystatins: defense proteins against phytophagous insects and Acari. **International journal of molecular sciences**, 17, n. 10, p. 1747, 2016.

MARTINS, C.; DE MARCO, R.; FARIAS, R.; MALGARIM, M. Noz pecã: panorama mundial. **Embrapa Clima Temperado–Artigo em periódico indexado (ALICE)**, 2019.

MCCOY, A. J.; GROSSE-KUNSTLEVE, R. W.; ADAMS, P. D.; WINN, M. D. *et al.* Phaser crystallographic software. **Journal of applied crystallography**, 40, n. 4, p. 658–674, 2007.

MCFERRIN, M. B.; SNELL, E. H. The development and application of a method to quantify the quality of cryoprotectant solutions using standard area–

detector X-ray images. **Journal of applied crystallography**, 35, n. 5, p. 538–545, 2002.

MESQUITA, C. M. *et al.* Manual do café: distúrbios fisiológicos, pragas e doenças do cafeeiro. **Belo Horizonte: EMATER**, 2016.

MOURA, G. T. D. Crystal structures and biophysical studies of phytocystatins. 2019.

MÜNTZ, K. Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seeds. **Journal of Experimental Botany**, 47, n. 5, p. 605–622, 1996.

NAGATA, K.; KUDO, N.; ABE, K.; ARAI, S. *et al.* Three-dimensional solution structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. japonica. **Biochemistry**, 39, n. 48, p. 14753–14760, 2000.

NAITO, Y.; FUJIE, M.; USAMI, S.; MUROOKA, Y. *et al.* The involvement of a cysteine proteinase in the nodule development in Chinese milk vetch infected with *Mesorhizobium huakuii* subsp. *rengei*. **Plant physiology**, 124, n. 3, p. 1087–1096, 2000.

NANDWANI, N.; SURANA, P.; NEGI, H.; MASCARENHAS, N. M. *et al.* A five-residue motif for the design of domain swapping in proteins. **Nature communications**, 10, n. 1, p. 1-13, 2019.

NELSON, D. L.; COX, M. M. Princípios de Bioquímica de Lehninger, 5ª edição. **São Paulo: Editora Artmed**, 6, 2011.

NEURATH, H. Proteolytic processing and physiological regulation. **Trends in biochemical sciences**, 14, n. 7, p. 268-271, 1989a.

NEURATH, H. The diversity of proteolytic enzymes. *Proteolytic enzymes: a practical approach*, p. 1-13, 1989b.

PARDE, Vinod D.; SHARMA, Hari C.; KACHOLE, Manvendra S. Protease inhibitors in wild relatives of pigeonpea against the cotton bollworm/legume pod borer, *Helicoverpa armigera*. **American Journal of Plant Sciences**, v. 3, p. 627-635, 2012.

PINA, A. S.; LOWE, C. R.; ROQUE, A. C. A. Challenges and opportunities in the purification of recombinant tagged proteins. **Biotechnology advances**, 32, n. 2, p. 366-381, 2014.

POLLEGIONI, P.; WOESTE, K. E.; CHIOCCHINI, F.; OLIMPIERI, I. *et al.* Landscape genetics of Persian walnut (*Juglans regia* L.) across its Asian range. **Tree Genetics & Genomes**, 10, n. 4, p. 1027–1043, 2014.

QUAIN, M. D.; MAKGOPA, M. E.; MÁRQUEZ-GARCÍA, B.; COMADIRA, G. *et al.* Ectopic phytocystatin expression leads to enhanced drought stress tolerance in soybean (*Glycine max*) and *A. rabidopsis thaliana* through effects on strigolactone pathways and can also result in improved seed traits. **Plant biotechnology journal**, 12, n. 7, p. 903–913, 2014.

RASHID, M. A.; KHATIB, F.; SATTAR, A. Protein preliminaries and structure prediction fundamentals for computer scientists. **arXiv preprint arXiv:1510.02775**, 2015.

RICKER, R.; SANDOVAL, L. Fast, reproducible size-exclusion chromatography of biological macromolecules. **Journal of chromatography A**, 743, n. 1, p. 43-50, 1996.

SENTHILKUMAR, R.; CHENG, C. P.; YEH, K. W. Genetically pyramiding protease-inhibitor genes for dual broad-spectrum resistance against insect and phytopathogens in transgenic tobacco. **Plant biotechnology journal**, 8, n. 1, p. 65–75, 2010.

SIEVERS, F.; WILM, A.; DINEEN, D.; GIBSON, T. J. *et al.* Fast, scalable

generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular systems biology**, 7, n. 1, p. 539, 2011.

STUBBS, M. T.; LABER, B.; BODE, W.; HUBER, R. *et al.* The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. **The EMBO journal**, 9, n. 6, p. 1939–1947, 1990.

SZEWIŃSKA, J.; SIMIŃSKA, J.; BIELAWSKI, W. The roles of cysteine proteases and phytocystatins in development and germination of cereal seeds. **Journal of plant physiology**, 207, p. 10–21, 2016.

TAN, Yanxiao *et al.* A phytocystatin gene from *Malus prunifolia* (Willd.) Borkh., MpCYS5, confers salt stress tolerance and functions in endoplasmic reticulum stress response in *Arabidopsis*. **Plant molecular biology reporter**, v. 34, n. 1, p. 62-75, 2016.

TAN, Yanxiao; LI, Mingjun; MA, Fengwang. Overexpression of MpCYS2, a phytocystatin gene from *Malus prunifolia* (Willd.) Borkh., confers drought tolerance and protects against oxidative stress in *Arabidopsis*. **Plant Cell, Tissue and Organ Culture (PCTOC)**, v. 123, n. 1, p. 15-27, 2015.

TOYOOKA, K.; OKAMOTO, T.; MINAMIKAWA, T. Mass transport of proform of a KDEL-tailed cysteine proteinase (SH-EP) to protein storage vacuoles by endoplasmic reticulum-derived vesicle is involved in protein mobilization in germinating seeds. **The Journal of cell biology**, 148, n. 3, p. 453–464, 2000.

VALADARES, N. F. *et al.* X-ray crystallography and NMR studies of domain-swapped canecystatin-1. **The FEBS journal**, v. 280, n. 4, p. 1028-1038, 2013. ISSN 1742-464X.

WAR, Abdul Rashid *et al.* Mechanisms of plant defense against insect herbivores. **Plant signaling & behavior**, v. 7, n. 10, p. 1306-1320, 2012.

WINN, M. D.; BALLARD, C. C.; COWTAN, K. D.; DODSON, E. J. *et al.* Overview of the CCP4 suite and current developments. **Acta Crystallographica Section D: Biological Crystallography**, 67, n. 4, p. 235–242, 2011.

WOLFENDEN, R. Degrees of difficulty of water-consuming reactions in the absence of enzymes. **Chemical reviews**, 106, n. 8, p. 3379–3396, 2006.

WOLFENDEN, R.; RADZICKA, A. A proficient enzyme. **Science**, 267, p. 90–93, 1995.

WOLFENDEN, R.; SNIDER, M. J. The depth of chemical time and the power of enzymes as catalysts. **Accounts of chemical research**, 34, n. 12, p. 938–945, 2001.

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

ZHU-SALZMAN, K. *et al.* Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine protease inhibitor. **Insect molecular biology**, v. 12, n. 2, p. 135-145, 2003.