



UnB

Instituto de Química

Programa de Pós-Graduação em Química

Dissertação de Mestrado

**CARACTERIZAÇÃO BIOQUÍMICA E BIOFÍSICA DA FRAÇÃO PROTEICA DO
FRUTO DA CAGAITEIRA
(*EUGENIA DYSENTERICA*)**

Marysol de Sousa Morais

ORIENTADORA: Profa.Dra. Andressa Regina Vasques Mendonça

Brasília, DF

2020



UnB

Instituto de Química

Programa de Pós-Graduação em Química

Dissertação de Mestrado

**CARACTERIZAÇÃO BIOQUÍMICA E BIOFÍSICA DA FRAÇÃO PROTEICA DO
FRUTO DA CAGAITEIRA
(*EUGENIA DYSENTERICA*)**

Marysol de Sousa Morais

Dissertação apresentada ao Programa de Pós-Graduação em Química da Universidade de Brasília, como parte dos requisitos para obtenção do título de mestra em química.

Orientadora: Profa. Dra. Andressa Regina Vasques Mendonça

Brasília, DF

2020

FOLHA DE APROVAÇÃO

Comunicamos a aprovação da Defesa de Dissertação do (a) aluno (a) **Marysol de Sousa Morais**, matrícula nº **17/0179117**, intitulada "*Caracterização bioquímica e biofísica da fração proteica do fruto da cagaiteira (Eugenia Dysenterica)*", apresentada no (a) Auditório Azul do Instituto de Química (IQ) da Universidade de Brasília (UnB) em 28 de janeiro de 2020.

Prof.^a Dra. Andressa Regina Vasques Mendonça
Presidente de Banca

Prof. Dr. Guilherme Dotto Brand
Membro Titular

Prof.^a Dra. Renata Vieira Bueno
Membro Titular IF/ USP

Prof. Dr. Caio de Oliveira Silva Gorgulho
Membro Suplente

Em 28 de janeiro de 2020.

Dedico esse trabalho a Deus, aos meus guias espirituais, a minha família e as minhas verdadeiras amizades.

“É muito melhor lançar-se em busca de conquistas grandiosas, mesmo expondo-se ao fracasso, do que alinhar se com os pobres de espírito, que nem gozam muito nem sofrem muito, porque vivem numa penumbra cinzenta, onde não conhecem nem vitória, nem derrota.”

Theodore Roosevelt

Agradecimentos

Agradeço a professora Sônia Maria de Freitas por ter me acolhido em seu laboratório e me dando assistência técnica e emocional quando necessário. Agradeço a Amanda por todo o apoio, ensinamentos e inspiração durante esse tempo que passei no laboratório de Biofísica .

Ao professor Osmino pelos ensinamentos, pela paciência e pelo apoio técnico e emocional. Agradeço aos estudantes Lucas Jeferson e Carlos por terem me ajudado na realização de alguns experimentos desse trabalho.

A todos os meus amigos do laboratório de Biofísica, Júlia, Gideane, Wallace, Paulo, Gustavo, Jonatas e os demais integrantes da família Biofísica pelos ensinamentos que vão muito além de conhecimentos acadêmicos. Agradeço pela companhia nas madrugadas e fins de semana, pelo apoio nas horas de desespero, pelas conversas sempre muito produtivas e por todo o conhecimento que compartilharam comigo. Serei eternamente grata por ter convivido com vocês e ter criado verdadeiros laços de amizade.

Ao Prof. Dr. Guilherme Dotto Brand por todos os conhecimentos compartilhados ao longo do mestrado e por praticamente ter me iniciado na área de estudos de proteínas.

Aos componentes da banca Prof. Dr. Guilherme Dotto Brand, Dra. Renata Vieira Bueno e Dr. Caio de Oliveira Gorgulho Silva por terem aceitado o convite para participar da banca da minha defesa de mestrado.

A professora Andressa Regina Vasques Mendonça por ter compartilhado esse projeto comigo.

Agradeço a minha família, as minhas amigas Graziela, Karine, Marina e Laisa, as minhas companheiras de república por tudo que foi compartilhado durante o período que moramos juntas, agradeço a Patrícia Figueiredo pelas noites de estudo compartilhada e por todo o apoio durante uma fase decisiva de construção desse trabalho.

Agradeço a minha família pelo apoio e compreensão, principalmente em relação aos momentos que precisei me ausentar por conta da vida acadêmica e a todos que me apoiaram de alguma maneira na realização desse trabalho e torceram para que as coisas dessem certo.

Resumo

Eugenia dysenterica é uma espécie nativa do Cerrado, conhecida popularmente como cagaíta ou cagaiteira. Essa espécie tem elevado potencial econômico, em função de suas propriedades e diversas aplicações, que vão desde o uso da madeira até o fruto. Mesmo frente ao amplo emprego, a quantidade de trabalhos referentes à sua composição química e a relação desta com suas propriedades ainda é pouco expressiva. O que mostra a necessidade e a importância da realização de estudos referentes a esta espécie, em particular, das suas propriedades farmacológicas, uma vez que, o uso de produtos naturais com propriedades medicinais é um nicho do mercado importante e em ascensão. Diante disso, o presente trabalho teve como objetivo: analisar a fração proteica do fruto da espécie *Eugenia dysenterica* por meio de análises qualitativa e quantitativa, baseadas em ensaios experimentais que englobam uma abordagem bioquímica e biofísica. Os resultados obtidos por meio da análise por eletroforese em gel de poli(acrilamida) e análise de massa – MALDI-TOF-TOF isolados da fração proteica revelaram a existência de três proteínas principais (cujas massas moleculares aproximadas obtidas foram 7 kDa, 9 kDa e 13 kDa) em ambas às metodologias empregadas. As proteínas citadas foram purificadas usando HPLC de fase reversa, a porção N-terminal da proteína de 7 kDa foi sequenciada e posteriormente foi feita a busca por similaridades no banco de dados utilizando o algoritmo blastp (NCBI). A sequência analisada apresentou identidade com 61 proteínas não específicas de transferência de lipídeos (nsLTP), sendo a maior identidade encontrada de 84,21% para *Syzygium oleosum* e *Eucalyptus grandis*. A proteína de 7 kDa foi classificada como sendo uma nsLTP putativa os dados das análises estruturais de dicroísmo circular mostraram que a proteína estudada é uma proteína termoestável o que corrobora com o fato da proteína ser uma nsLTP putativa.

Palavras-chaves: *eugenia dysenterica*, transferência de lipídeos, laxativo

Abstract

Eugenia dysenterica is a native species of the cerrado, popularly known as cagaita or cagaiteira, which has high economic potential, due to its properties and its various applications, ranging from wood use to fruit consumption. Despite its wide use, the amount of work related to its chemical composition and its relation with its properties is still not expressive, which shows the necessity and importance of conducting studies regarding this species, in particular, its pharmacological properties, as the use of natural products with medicinal properties is an important and growing niche market. Therefore, the present study aimed to: analyze the protein fraction of *eugenia dysenterica* fruit through a qualitative and quantitative analysis based on experimental assays that include a biochemical and biophysical approach. The results obtained through the characterization) of the isolated components of the protein fraction revealed the existence of three proteins whose approximate molecular masses obtained were 7 kDa, 9 kDa and 13 kDa) in both methodologies employed (analysis by polyacrylamide gel electrophoresis and mass analysis (MALDI-TOF-TOF). Cited proteins were purified using reverse phase HPLC. Partial N-terminal sequencing of the 7 kDa protein was performed and the search for similarities in the database using the blastp algorithm (NCBI). The sequence analyzed presented identity with 61 non-specific lipid transfer proteins (nsLTP), being the highest identity found of 84.21% for *syzygium oleosum* and *eucalyptus grandis*. The protein studied was classified as a putative nsLTP. Data from the circular dichroism structural analysis showed that the protein studied is a thermostable protein which corroborates the fact that the protein is a putative nsLTP.

Keywords: *eugenia dysenterica*, Lipids transfer, laxative

Lista de Aminoácidos

Aminoácido	Código de três letras	Código de uma letra
Alanina	Ala	A
Arginina	Arg	R
Asparagina	Asn	N
Ácido aspártico	Asp	D
Cisteína	Cis	C
Ácido glutâmico	Glu	E
Glutamina	Gln	Q
Glicina	Gly	G
Histidina	His	H
Isoleucina	Ile	I
Leucina	Leu	L
Lisina	Lys	K
Metionina	Met	M
Fenilalanina	Phe	F
Prolina	Pro	P
Serina	Ser	S
Treonina	Thr	T
Triptofano	Trp	W
Tirosina	Tyr	Y
Valina	Val	V

Lista de abreviaturas e acrônimos

AMPs : Peptídeos antimicrobianos (AntiMicrobial Peptides)

BCA: Ácido-Bicinconínico (Bicinchoninic Acid)

BLAST: Ferramenta básica de pesquisa de alinhamento local (*Basic Local Alignment Search Tool*)

BSA: Albumina soro bovina (Bovine Serum Albumin)

CD :Dicroísmo Circular (Circular Dichroism)

CDD : Banco de domínios conservados (Conserved Domain Database)

CG: Cromatografia gasosa
de SDS

DTT DL: Dithiothreitol
Electrophoresis)

ESI: Ionização por electrospray (Electron Spray Ionization)

HPLC: Cromatografia Líquida de Alta Eficiência (*High Performance Liquid Chromatography*).
Laser Desorption Ionization)

LTPs: Proteínas transportadoras de lipídeos

MALDI Ionização/Dessorção de Matriz Assistida por Laser (*Matrix Assisted*
National Center for Biotechnology Information)

nsLTPs proteínas não específicas de transferência de lipídeos

PAGE :Eletroforese em gel de poliacrilamida (*PolyAcrylamide Gel*)

PR: Proteínas Relacionadas à patogênese (*Pathogenesis Related*)

PSA: Persulfato de amônia

PTH: Feniltiohidantoína

SDS: Dodecilsulfato de sódio (*Sodium Dodecyl Sulfate*)

SDS-PAGE: Eletroforese desnaturante em gel de poliacrilamida na presença

TEMED: Tetrametiletlenodiamina

TFA: Ácido trifluoracético (*TriFluoroacetic Acid*)

TOF: Tempo de voo (*Time-Of-Flight*)

TRIS: Tris (Hidroximetil) aminometano

Lista de Tabelas

Tabela 1 - Frações contendo coletadas no ensaio de cromatografia HPLC que apresentava as proteínas de interesse e suas respectivas massas	27
Tabela 2 - Rendimentos extração e purificações	31
Tabela 3 - Conteúdo de estrutura secundária da proteína de 7 kDa na polpa de <i>E. dysenterica</i> em diferentes pHs obtidos por deconvolução dos espectros de DC utilizando o programa CDNN	35
Tabela 4 - Conteúdo de estrutura secundária da proteína de 9 kDa presente na polpa de <i>E. dysenterica</i> obtido por deconvolução dos espectros de DC por meio do programa CDNN. ...	38
Tabela 5 - Conteúdo de estrutura secundária da proteína de 13 kD presente na polpa de <i>E. dysenterica</i> obtido por deconvolução dos espectros de DC por meio do programa CDNN. ...	42

Lista de Figuras

Figura 1 – Árvore da Cagaiteira ^[14]	4
Figura 2 - Tronco da cagaiteira ^[14]	4
Figura 3 - Flores da cagaiteira ^[14]	4
Figura 4 - Fruto jovem (A) e fruto maduro cortado ao meio (B) ^[15]	5
Figura 5 - Alguns compostos químicos encontrados em diversas partes de <i>Eugenia dysenterica</i> ^[17]	7
Figura 6 - Níveis de organização das proteínas ^[23]	8
Figura 7 - Perfil da ACPase II deglicosilada SDS PAGE 12% apresentando (MM) Marcador Molecular (Thermo Scientific); (linha 1) ACPase II intacta $94,4 \pm 4,0$ kDa (87 μ g); (linha 2) ACPase II pós deglicosilação $67 \pm 5,0$ kDa (100 μ g). Massas moleculares foram determinadas através do programa Image Lab 5 (Bio-Rad) ^[29]	9
Figura 8 - Mecanismo de degradação de Edman, Fenilisotiocianato reage com um grupo amino-terminal (A), formando um derivado cíclico do feniltiocarbamoil (B), que em condições ácidas é clivado dando origem a um derivado da tiazolinona(C) que é posteriormente tratado com ácido para dar origem a um derivado mais estável(D) (figura retirada e adaptada da literatura) ^[35]	11
Figura 9 - Esquema simplificado de um espectrômetro de massa contendo seus principais componentes: fonte de íons, analisador m/z, detector, sistema de vácuo e o computador.....	12
Figura 10 - Espectros de CD típicos de estruturas do tipo α -hélices, folhas- β e estruturas irregulares ^[24]	14
Figura 11 - Esquema da deconvolução espectral da albumina sérica humana ^[24]	15
Figura 12 - Fluxograma da metodologia realizada para isolamento e caracterização das proteínas presentes na fração proteica do fruto da cagaiteira.....	18
Figura 13 - Curva de calibração construída com soluções padrão de albumina sérica bovina para a quantificação de proteínas totais no material proveniente da extração da fração proteica do fruto da cagaiteira.	23
Figura 14 - Perfil da fração proteica em gel de tris-tricina corado com prata usando o marcador de massa molecular <i>SeeBlue™ Plus2 Pre-stained Protein Standard</i> com uma mistura de 10 proteínas com massa entre 3 e 198 kDa. Quatro bandas proteicas foram observadas, uma proteína de aproximadamente 7kDa (A), uma banda na altura de 9 kDa (B), uma proteína entre 13 e 14 kDa(C) e por fim pode ser observado uma banda de aproximadamente 20 kDa (D)..	24
Figura 15 - Perfil da fração proteica em gel tris-glicina corado com comassie blue e usando o marcador de massa molecular <i>SeeBlue™ Plus2 Pre-stained Protein Standard</i> com uma mistura de 10 proteínas com massa entre 3 e 198 kDa.(MM) As 4 bandas correspondentes a uma proteína de aproximadamente 7 kDa (A), uma banda na altura de 9 kDa (B) estão evidenciadas. A seta amarela aponta uma proteína entre 13 e 14 kDa(C) e por fim pode ser observado uma banda de aproximadamente 20 kDa (D).....	25
Figura 16 - Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico. Extrato aquoso do fruto de <i>Eugenia dysenterica</i> , observa-se os componentes de massas aproximadas de 7,9 e 13kDa e o dupla carga do componente de massa aproximada de 13kDa.....	26
Figura 17 - Corrida de triagem da fração proteica da cagaita foi usada uma coluna de fase reversa C18 (shimadzu VP ODS) e as amostras foram eluídas com gradiente de	

0 a 60 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 50 minutos.....	27
Figura 18 - Perfil cromatográfico das corridas feitas para isolamento das frações de interesse realizadas em uma coluna de fase reversa C18 (shimadzu VP ODS) eluidas com gradiente de 0 a 60 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 50 minutos.	28
Figura 19 - Cromatograma representativo das recromatografias das frações 15.1 e 16 correspondente a proteína de 7,284 kDa as amostras foram eluidas em uma coluna de fase reversa C18 (shimadzu VP ODS) com gradiente de 0 a 30 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 25 minutos a fração de interesse coletada está destacada com uma seta.....	29
Figura 20 - Cromatograma representativo das recromatografias das frações 18 correspondentes a proteína de 9,252kDa, as amostras foram eluidas em uma coluna de fase reversa C18 (shimadzu VP ODS) com gradiente de 0 a 35 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 20 minutos a fração de interesse coletada está destacada com uma seta.	29
Figura 21 - Cromatograma representativo das recromatografias das frações 19 e 23 correspondente a proteína de 13,7 kDa as amostras foram eluidas em uma coluna de fase reversa C18 (shimadzu VP ODS) com gradiente de 0 a 42 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 20 minutos a fração de interesse coletada está destacada com uma seta.....	30
Figura 22 - Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico,para a fração correspondente a proteína com massa aproximada de 7kDa. ..	30
Figura 23 - Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico,para a fração correspondente a proteína com massa aproximada de 9kDa. ..	30
Figura 24 - Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico,para a fração correspondente a proteína com massa aproximada de 13kDa.....	31
Figura 25 - Sequência parcial da proteína de 7,2 kDa obtida por degradação automática de Edman onde X são prováveis cisteínas.....	32
Figura 26 - Alinhamento múltiplo da sequência primária parcial da proteína de <i>E. dysenterica</i> de 7kDa com 10 nsLTP's homólogas disponíveis no banco de sequências de proteínas não redundantes (nr). As proteínas estão identificadas com o número de acesso. Os resíduos conservados de Cys estão marcados de azul. Os demais resíduos conservados estão marcados de verde e os resíduos semi-conservados estão marcados de cinza.	32
Figura 27 - Representação dos domínios conservados da sequência obtida após comparação com a à base de dados CDD (<i>Conserved Domain Database</i>), que demonstra um único domínio conservado localizado entre os resíduos Ser4 e Gly38, sendo este pertencente a subfamília das proteínas não específicas de transferência de lipídeos do tipo 2 (nsLTP2).....	33
Figura 28 - Espectros dicróicos da proteína de 7 kDa na região distante da luz UV (190-260 nm) nos pH 4,0; 7,0 e 9,0 obtidos a 25 °C. A concentração proteica foi 0,19 mg/mL em solução nos tampões acetato de sódio 5 mM para o pH 4,0 e Tris HCl 5 mM para os pHs 7,0 e 9,0...	35
Figura 29 - Curvas de termoestabilidade e espectros dicróicos da proteína de 7 kDa. (A) e espectros dicróicos da proteína de 7 kDa presente na polpa de <i>E. dysenterica</i> nos pHs 4,0 (B), 7,0 (C) e 9,0 (D).Curvas de desnaturação térmica foram obtidas no comprimento de onda fixado	

em 208 nm (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul), com variação da temperatura de 25 °C á 95 °C. Os espectros dicroicos da proteína foram obtidos na região distante da luz-UV (198-260 nm), de 25 °C a 95 °C, com intervalos de 10 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul)..... 37

Figura 30 - Espectros dicroicos da proteína de 9 kDa na região distante da luz UV (198-260 nm) nos pHs 4,0; 7,0 e 9,0 com a concentração proteica de 0,27 mg/mL em solução nos tampões acetato de sódio 5 mM para o pH 4,0 e Tris HCl 5 mM para os pHs 7,0 e 9,0. Este resultado foi observado em todos os pHs testados, porém, nos pHs 7,0 e 9,0, o percentual de estrutura secundária é maior para alfa hélice (13,5%), em comparação ao pH 4,0 que é de 11,6% (**Tabela 3**). Nos pHs 7,0 e 9,0, o conteúdo de folha beta paralela reduz em ~2,5% Esses resultados mostram que a proteína de 9 kDa apresenta-se mais estruturada na faixa neutra e básica. 37

Figura 31 - Curvas de termo estabilidade e espectros dicroicos da proteína de 9 kDa. Curvas de termo estabilidade (A) e espectros dicroicos da proteína de 9 kDa presente na polpa de *E. dysenterica* nos pHs 4,0 (B), 7,0 (C) e 9,0 (D). As curvas de desnaturação térmica foram obtidas no comprimento de onda fixado em 208 nm, com variação da temperatura de 25 °C para 95 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul). Os espectros dicroicos da proteína foram obtidos na região distante da luz-UV (198-260 nm), de 25 °C a 95 °C, com intervalos de 10 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul). 39

Figura 32 - Espectros dicroicos de desnaturação e de renaturação da proteína de 9 kDa *E. dysenterica* na região distante da luz UV nos pHs 4,0, pH 7,0 e pH 9,0 obtidos a 25 °C (linhas contínuas) e após ensaio de desnaturação térmica (25 a 95 °C) com retorno a 25 °C (linhas pontilhadas). Os espectros dicroicos da proteína de 9 kDa foram obtidos de 198-260 nm antes e após desnaturação térmica. 40

Figura 33 - Espectros dicroicos da proteína de 13 kDa na região distante da luz UV nos pHs 4,0; 7,0 e 9,0 obtidos a 25 °C. A concentração proteica foi 0,27 mg/mL em solução nos tampões acetato de sódio 5 mM para o pH 4,0 e Tris HCl 5 mM para os pHs 7,0 e 9,0. 41

Figura 34 - Curvas de termoestabilidade e espectros dicroicos da proteína de 13 kDa em diferentes pH's. Curvas de termo estabilidade (A) e espectros dicroicos da proteína de 13 kDa presente na polpa de *E. dysenterica* nos pHs 4,0 (B), 7,0 (C) e 9,0 (D). As curvas de desnaturação térmica foram obtidas no comprimento de onda fixado em 218 nm, com variação da temperatura de 25 °C para 95 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul). Os espectros dicroicos da proteína foram obtidos na região distante da luz-UV de 25 °C a 95 °C, com intervalos de 10 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul). 43

Figura 35 - Espectros dicroicos da proteína de 13 kDa durante e após ensaio de desnaturação. Espectros dicroicos da proteína de 13 kDa presente na polpa de *E. dysenterica* na região distante da luz UV nos pHs 4,0, pH 7,0 e pH 9,0 obtidos a 25 °C (linhas contínuas) e após ensaio de desnaturação térmica (25 a 95 °C) com retorno a 25 °C (linhas pontilhadas). Os espectros dicroicos da proteína de 9 kDa foram obtidos de 198-260 nm antes e após desnaturação térmica. 44

Sumário

INTRODUÇÃO E OBJETIVOS	1
1 REVISÃO BIBLIOGRÁFICA.....	3
1.1 Descrição da <i>Eugenia dysenterica</i>	3
1.2 Aproveitamento de <i>Eugenia dysenterica</i>	5
1.3 Aspectos nutricionais e constituintes químicos da cagaiteira.....	6
1.4 Estudo de proteínas.....	7
1.4.1 Degradação de Edman	10
1.4.2 Espectrometria de massas aplicada ao estudo de proteínas	12
1.4.3 Alinhamento de sequências para o estudo de proteínas.....	12
1.4.4 Dicroísmo circular	13
1.4.5 Proteínas em plantas	15
2 MATERIAIS E MÉTODOS	17
2.1 Extração da fração proteica	18
2.2 Quantificação da fração proteica	19
2.3 Análise da fração proteica	19
2.4 Purificação da fração proteica	20
2.5 Sequenciamento por Edman e análise da sequência obtida.....	21
3 RESULTADOS E DISCUSSÕES	23
3.1 Extração e quantificação da fração proteica	23
3.2 Análise da fração proteica por eletroforese em gel de poliacrilamida	23
3.3 Purificação da fração proteica	26
3.4 Sequenciamento por degradação de Edman	31
4 CONCLUSÃO	45
5 PERSPECTIVAS	46
6 REFERÊNCIAS	47

INTRODUÇÃO E OBJETIVOS

Os frutos provenientes das espécies nativas do cerrado são conhecidos principalmente por suas características sensoriais: cor, sabor e odor que são bem peculiares, além de apresentarem elevado valor nutricional. No entanto, apesar de apresentarem estas características, o potencial comercial dessas espécies ainda é pouco explorado ^[1].

Dentre as espécies nativas do Cerrado que apresentam elevado potencial comercial está a cagaiteira (*Eugenia dysenterica*). Essa árvore frutífera de porte médio apresenta diversas aplicações, desde o seu tronco até seus frutos, possibilitando um aproveitamento integral da planta ^[2].

Entre as possíveis aplicações da cagaiteira, as medicinais são bem difundidas entre os moradores do Cerrado, utilizando principalmente as suas folhas e frutos. Suas folhas podem ser usadas na forma de extrato aquoso (chás) para o tratamento de diversos problemas de saúde, dentre eles, icterícia, diarreia e problemas de coração. Já seus frutos, que também apresentam propriedades medicinais, segundo o conhecimento popular, quando maduras têm ação laxativa, além de serem muito empregados na culinária ^[3,4, 5,6].

Entretanto, mesmo que a cagaiteira seja uma planta com diversas aplicações e propriedades notáveis, a quantidade de estudos científicos sobre as propriedades medicinais e constituintes químicos e biológicos dos frutos desta planta não são abundantes. Grande parte do conhecimento sobre essa espécie é baseado no senso comum. E a maioria dos estudos sobre essa espécie é relacionado às folhas e seus constituintes químicos. Em contrapartida, estudos relacionados aos constituintes químicos dos frutos são escassos ^[2].

Portanto, são necessários, estudos de natureza bioquímica e biofísica referentes aos constituintes químicos do fruto da cagaiteira, com o objetivo de caracterizar as biomoléculas, favorecendo o conhecimento de suas propriedades e possíveis aplicações, sobretudo na área alimentícia e farmacológica ^[2].

Diante do exposto, o objetivo geral desse trabalho consiste em caracterizar as proteínas presentes no fruto de *Eugenia dysenterica* por ensaios bioquímicos e biofísicos afim de delinear condições para seu armazenamento, processamento e comercialização para um maior aproveitamento de seu potencial econômico. Para tal, os seguintes objetivos específicos devem ser contemplados:

- Obtenção do extrato bruto dos frutos congelados de *Eugenia dysenterica*.
- Extração da fração proteica a partir do extrato bruto obtido dos frutos congelados.
 - Purificação das proteínas presentes na fração proteica extraída do referido fruto por meio de HPLC Cromatografia Líquida de Alta Eficiência (*High Performance Liquid Chromatography*).
 - Caracterização da estrutura primária das proteínas identificadas na fração proteica do fruto da cagaiteira por meio de sequenciamento de Edman.
 - Caracterização da estrutura secundária das proteínas presentes no fruto da cagaiteira por dicroísmo circular.

1 REVISÃO BIBLIOGRÁFICA

1.1 Descrição da *Eugenia dysenterica*

A espécie *Eugenia dysenterica* é conhecida popularmente como cagaita ou cagaiteira devido ao efeito laxativo atribuído popularmente aos frutos da espécie quando esses estão maduros e são ingeridos em excesso, ou quando estes estão fermentados ^[1,6]. A cagaiteira faz parte das mais de 50.000 mil espécies conhecidas que são nativas do cerrado podendo ser encontrada tanto no Cerradão Mesotrófico, quanto no Cerradão Distrófico estando assim adaptada a solos pobres em nutrientes ^[8]. A cagaita pode ser encontrada na Bahia, Distrito Federal, Goiás, Maranhão, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Pará, Piauí, São Paulo e Tocantins ^[9].

A cagaiteira pertence à família *Myrtaceae* e faz parte de um dos maiores gêneros dessa família, o gênero *Eugenia*. É uma árvore frutífera de porte médio e crescimento lento que quando adulta pode atingir até 15 metros de altura (**Figura 1**). Possui troncos e ramos tortuosos e casca grossa com fissuras (**Figura 2**). Sua copa é alongada e densa, com ramos quadrangulares e glabros quando adultos (**Figura 1**). Os botões, pedicelos flores e ramos quando jovens são pubérulos. Suas folhas são membranáceas opostas, ovado - oblongas, simples, curto-pecioladas a subsésseis, glabras, aromáticas e caducas na floração ^[10].

A floração ocorre anualmente entre agosto e setembro, geralmente acompanhando as primeiras chuvas e não costuma durar mais que uma semana. Suas flores são axilares, solitárias ou organizadas em arranjos de três (**Figura 3**)^[11]. A floração, a produção de nova folhagem e a frutificação se dão no espaço de um mês e seus principais polinizadores são abelhas ^[12].

Os frutos da cagaiteira são bagas globulares, verdes quando imaturos (**Figura 4 A**) e amarelo pálido quando maduros (**Figura 4B**). A maturação desses frutos é rápida e coincide com o período chuvoso compreendendo um período de aproximadamente 4 semanas entre a formação do fruto e sua completa maturação^[8, 13].

Os frutos maduros possuem de 2 a 3 cm de diâmetro com peso entre 15 e 20 g e possuem de duas a três sementes por fruto, que são envoltas por uma polpa levemente ácida. Apresenta um cálice seco aderido ao fruto. Possuem mesocarpo e endocarpo suculentos, e casca

membranácea brilhante^[8, 13]. Suas sementes são de formato oval, achatado ou elipsoide, medindo entre 0,8 cm e 2 cm ^[8, 10].



Figura 1 – Árvore da Cagaiteira ^[14].



Figura 2 - Tronco da cagaiteira ^[14].



Figura 3 - Flores da cagaiteira ^[14].



Figura 4- Fruto jovem (A) e fruto maduro cortado ao meio (B)^[15].

1.2 Aproveitamento de *Eugenia dysenterica*

A cagaiteira é uma espécie com elevado interesse econômico devido às diversas aplicações apresentadas por essa planta. A casca, além de servir à indústria de curtume, é utilizada na medicina popular como antidiarreica. Por apresentar uma quantidade de súber considerável, e ter uma espessura que varia de 1,0 a 2,0 cm, ela é muito empregada também na indústria de cortiça ^[16]. A madeira do caule da cagaiteira é densa, com densidade de 0,82 g/cm², dura e de textura fina, porém, por possuir uma baixa qualidade, é comumente usada como mourão, lenha e carvão ^[17].

Suas folhas podem ser usadas como excelentes pastos arbóreos, e, além disso, apresentam variadas propriedades medicinais, sendo utilizadas na medicina popular como antidiarreicas, para problemas do coração, e no tratamento de diabetes e icterícia^[19,20].

Os frutos são os maiores responsáveis pelo interesse econômico dessa espécie, pois apresentam diversas aplicações culinárias. Além de serem muito saborosos para consumo *in natura*, os frutos servem de base para diversas receitas como geleia, sorvetes, compotas, *chutneys* e diversas bebidas^[18].

A perecibilidade dos frutos prejudica a comercialização deles. O fruto perece em apenas três dias se conservado a 28 °C, mas pode ser conservado por até treze dias quando armazenado em geladeira a 15 °C. Sua polpa se for congelada, pode ser conservada em condições de consumo por mais de um ano, sem perder o sabor ^[21].

Apesar de serem saborosos, seus frutos devem ser consumidos com parcimônia pois se consumidos em excesso, quando maduros ou fermentados, têm efeito laxativo,Essa propriedade laxativa foi comprovada com estudos *in vivo* em ratos e atribuída a fração proteica do fruto [22].O efeito laxativo juntamente com a precibibilidade dos frutos pode ser apontado como um dos fatores que impede a maior difusão do consumo desses frutos.

1.2 Aspectos nutricionais e constituintes químicos da cagaiteira.

Em 100 g do fruto da cagaiteira tem-se em média 20 kcal, cerca 0,82 g de proteínas, 0,44 g de lipídios, 3,08 g de carboidratos, 1,04 g de fibras alimentares e 94% de umidade [2]. A cagaiteira é considerada boa fonte de vitamina C (18-72 mg/100 g), vitamina B2 (0,4 mg/100 g), cálcio (172,8 mg/100 g), magnésio (62,9 mg/100 g) e ferro (3,9 mg/100 g)^[9].

Os constituintes de diferentes partes da cagaiteira têm sido identificados por diversas técnicas. No extrato etanólico das folhas são identificados taninos, flavonoides, terpenos e saponinas por meio da técnica analítica de cromatografia de camada delgada [15]. A análise por *High Performance Liquid Chromatography* (HPLC), em português, Cromatografia Líquida de Alta Eficiência detectou a presença de polifenóis, taninos e flavonoides nos extratos etanólico e aquoso [16,17].

Na polpa, por meio da técnica de micro extração em fase sólida acoplada à cromatografia gasosa (CG) com detector de massas, foram identificadas algumas substâncias voláteis. Os ésteres hexanoato de etila e butanoato de etila foram identificados como os componentes majoritários [18].

No óleo essencial das folhas obtido pelo processo de hidro destilação foram identificados compostos pertencentes ao grupo dos sesquiterpenos [19].No óleo essencial que é obtido dos frutos semimaduros e imaturos, são identificados α -limoneno, (E)- β -cimeno e β -pineno. Nos frutos maduros são identificados γ -muuroleno , β -cariofileno e α -humuleno [19].Já no óleo essencial obtido de sementes foram identificados os sesquiterpenos β -cariofileno , δ -cadineno e β -pineno, entre outros [19].

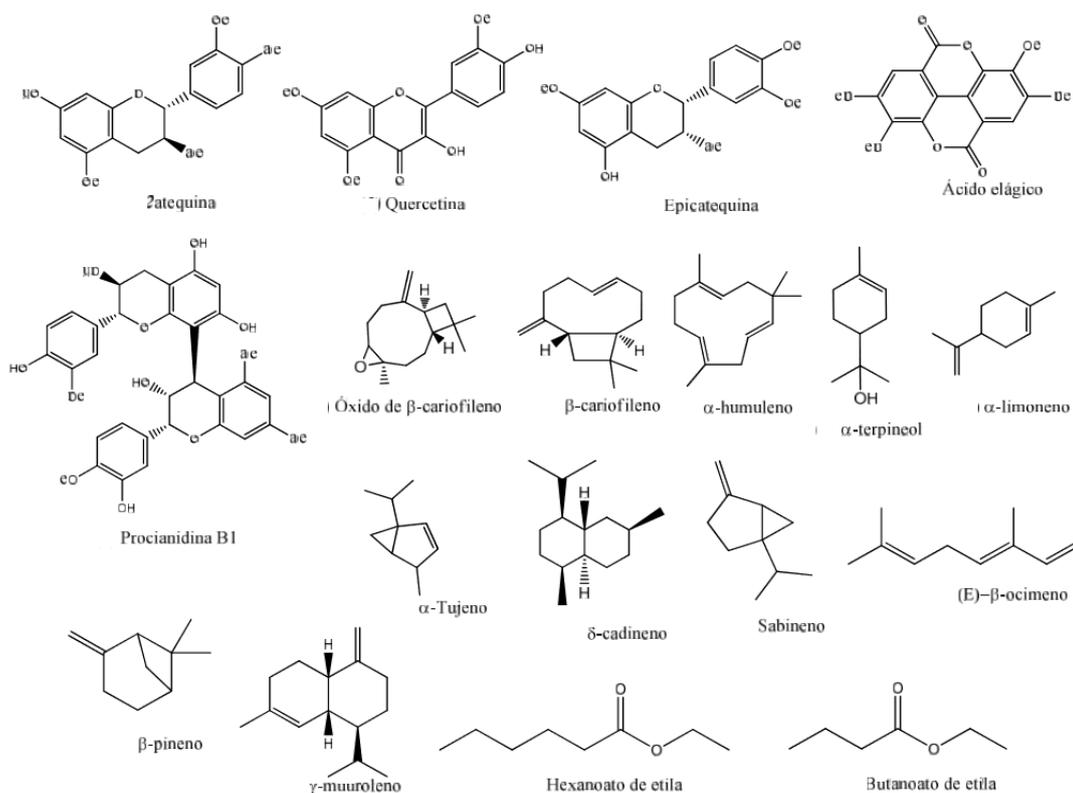


Figura 5 - Alguns compostos químicos encontrados em diversas partes de *Eugenia dysenterica* [17].

1.1 Estudo de proteínas

As proteínas são macromoléculas que são encontradas de maneira abundante nas células dos organismos vivos. Essas macromoléculas são compostas pela combinação de 20 aminoácidos que se ligam em diversas ordens, em seqüências lineares por meio de ligações covalentes conhecidas como ligações peptídicas [23,24].

Essa grande variedade de seqüências que podem ser formadas pela combinação dos aminoácidos, confere às proteínas uma grande importância biológica, pois a partir da combinação desses blocos formadores os organismos podem sintetizar uma gama de produtos diferentes entre si, dentre eles: enzimas, hormônios, anticorpos, antibióticos, venenos, dentre outros milhares de compostos biologicamente importantes [23,24,25].

As proteínas possuem diferentes níveis de organização ordenados por níveis de complexidade. Comumente quatro níveis são descritos: (a) Estrutura primária que consiste na seqüência de resíduos de aminoácidos que compõem a cadeia polipeptídica, (b) Estrutura

secundária que é o arranjo espacial formado pelo na interação dos aminoácidos próximos da cadeia polipeptídica que se organizam em torno de um eixo imaginário. Essas estruturas são estabilizadas por interações de hidrogênio e as conformações mais estudadas são a alfa-hélice e a estrutura folha beta-plana. (c) Estrutura terciária que descreve a estrutura tridimensional das proteínas mantidas por meio de interações não covalentes. (d) Estrutura quaternária é o arranjo formado por proteínas que possuem mais de uma subunidade polipeptídica [23,24,26].

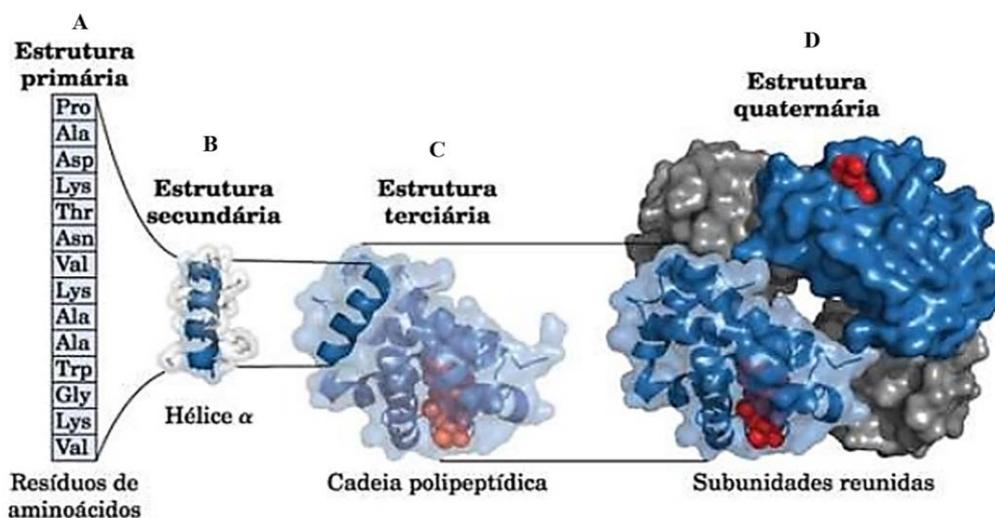


Figura 6 - Níveis de organização das proteínas [23].

Devido à grande diversidade e importância biológica, o estudo das proteínas é uma área de relevância. Para conhecer uma proteína é necessário que esta seja isolada para posteriormente serem usadas técnicas e equipamentos que possibilitem analisar as suas propriedades. Para realizar a extração da fração proteica de interesse, uma técnica muito utilizada é a *Salting out*. Essa técnica consiste na precipitação de proteínas por meio de adição de sais que diminuem a solubilidade das proteínas no meio. O sal mais usado neste tipo de extração é o sulfato de amônio $(\text{NH}_4)_2 \text{SO}_4$ devido à sua elevada solubilidade em água [23,25]. Essa é uma técnica clássica de extração que é usada até hoje para extração da fração proteica de extratos vegetais dentre elas a extração de BTCl [27,28] e da fração proteica fruto da cagaiteira [21].

A eletroforese é uma técnica comumente usada para separação de proteínas, essa técnica é baseada na migração de proteínas carregadas em um campo elétrico. Atualmente, esse método não é usado para a separação de grandes quantidades de proteína, pois já existem métodos mais eficazes e modernos. Mesmo assim, ainda hoje esse método é usado para análise preliminar das proteínas, servindo para confirmar a presença das proteínas de interesse e estimar seu peso

molecular na amostra estudada [23,24,25]. Na (**Figura 7**) um exemplo da aplicação da técnica de eletroforese para estudo do Perfil da ACPase II deglicosilada. [29]

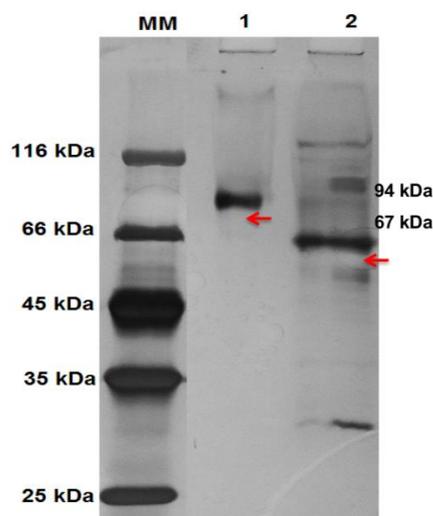


Figura 7 - Perfil da ACPase II deglicosilada SDS PAGE 12% apresentando (MM) Marcador Molecular (Thermo Scientific); (linha 1) ACPase II intacta $94,4 \pm 4,0$ kDa ($87 \mu\text{g}$); (linha 2) ACPase II pós deglicosilação $67 \pm 5,0$ kDa ($100 \mu\text{g}$). Massas moleculares foram determinadas através do programa Image Lab 5 (Bio-Rad)^[29].

Após uma análise preliminar o próximo passo é a purificação da(s) proteína(s) de interesse. A purificação das proteínas é feita com base nas suas propriedades, sendo os métodos cromatográficos os mais usados para esse fim. Esses métodos consistem basicamente na separação de componentes pela eluição da fração proteica com um solvente (fase móvel) através de uma coluna (fase estacionária) [23,25].

A separação das proteínas se dá por sua afinidade com a fase móvel ou com a fase estacionária, o que determina o tempo em que a fração de interesse será eluída. O procedimento clássico consiste no preenchimento de colunas de vidro com sólidos porosos e quimicamente adequados, por onde a fração proteica será eluída com um solvente adequado (fase móvel) [23].

A *High Performance Liquid Chromatography* (HPLC) é um método mais refinado e moderno dentro dessa técnica. Nesse equipamento bombas de alta pressão aumentam a velocidade de eluição das proteínas pela coluna, o que diminui o tempo de análise. Além disso, nesse equipamento as colunas são projetadas e construídas de forma que em um pequeno espaço se tenha uma grande área superficial, o que aumenta a efetividade do método [23]. A purificação de peptídeos antimicrobianos presentes na secreção cutânea do anuro *Hypsiboas raniceps* [30] é um exemplo do uso dessa técnica.

Após a purificação da proteína, o passo seguinte deve ser a elucidação da sua estrutura primária. As técnicas mais usadas para tal objetivo são: Degradação química de Edman que é uma técnica clássica que consiste no sequenciamento por meio da degradação seletiva dos resíduos amino terminal de um peptídeo preservando as demais ligações peptídicas [21]. E Métodos baseados na técnica de espectroscopia de massas, dentre eles: *Matrix-Assisted Laser Desorption/Ionization TOF* (MALDI- TOF) e o Método MS sequencial [23]. O sequenciamento de um peptídeo aniônico antimicrobiano da secreção cutânea de *Osteocephalus taurinus* foi feito usando o método MALDI-TOF-MS/MS [31].

Com a estrutura primária da proteína de interesse elucidada ou até mesmo parte dela pode então ser feito um estudo baseados em técnicas de bioinformática que podem levar a informações importantes sobre a proteína estudada. As estruturas secundárias de proteínas são o nível de organização imediatamente posterior as sequencias de aminoácidos que compõem as cadeias peptídicas. Para predição dessas estruturas umas das técnicas que podem ser utilizadas é a de Dicroísmo circular [23,24]. Essa técnica permite o estudo da estrutura secundária e termo estabilidade de proteínas, tais como os feitos nos estudos de uma α 2- macroglobulina bacteriana [32] e da serino protease transmembrana II humana [33].

1.4.1 Degradação de Edman

A degradação de Edman é uma metodologia usada para sequenciamento de aminoácidos de um peptídeo desenvolvida por Pehr Edman. Esse método consiste na clivagem individual de um resíduo amino-terminal sem afetar as ligações peptídicas entre outros resíduos de aminoácidos da cadeia do polipeptídeo. Essa técnica pode ser repetida mais uma vez para identificar o aminoácido seguinte da cadeia [34].

A grande limitação dessa técnica está relacionada ao rendimento que só é satisfatório para aproximadamente 50 aminoácidos. O número de resíduos de aminoácidos identificados é limitado, devido à derivação cíclica que nem sempre poder ser completada, e, além disso, não é possível fazer o sequenciamento de peptídeos que tenham algum tipo de bloqueio na porção N-terminal [34].

Na degradação de Edman, em condições levemente alcalinas o Fenilisotiocianato reage com um grupo amino-terminal, formando um derivado cíclico do feniltiocarbamoil, que em condições ácidas é clivado dando origem a um derivado da tiozalinona que é tratado em meio ácido para dar origem a um derivado da feniltiohidantoína (PTH) que é mais estável [35].

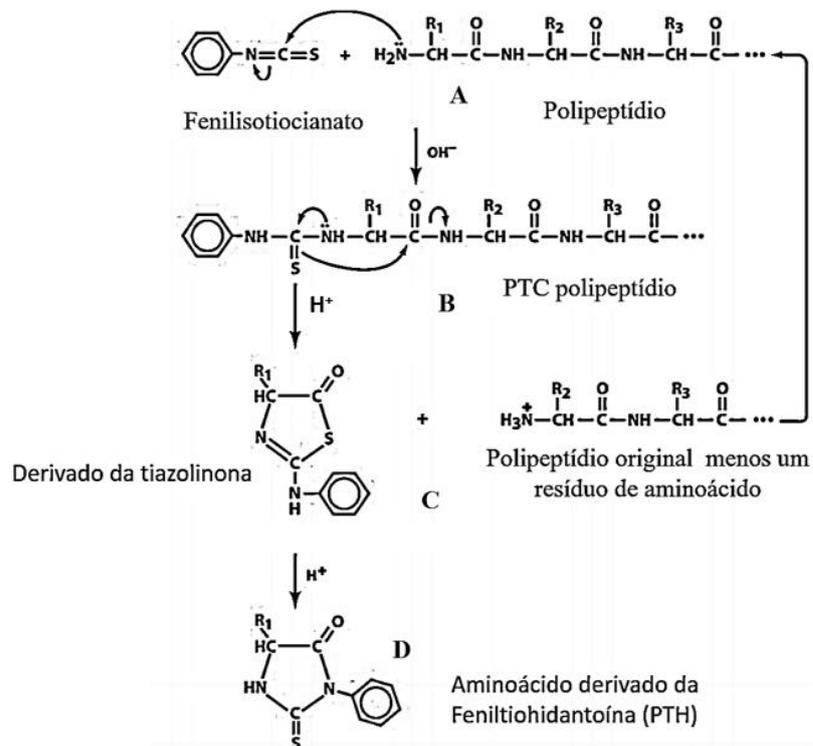


Figura 8 - Mecanismo de degradação de Edman, Fenilisotiocianato reage com um grupo amino-terminal (A), formando um derivado cíclico do feniltiocarbamoil (B), que em condições ácidas é clivado dando origem a um derivado da tiazolinona(C) que é posteriormente tratado com ácido para dar origem a um derivado mais estável(D) (figura retirada e adaptada da literatura) ^[35].

1.4.2 Espectrometria de massas aplicada ao estudo de proteínas

A Espectrometria de Massas (EM) é uma técnica que pode ser usada para se determinar quantitativamente e qualitativamente a composição de uma amostra, e pode ser utilizada também para a elucidação de estruturas de moléculas. Essa técnica é baseada na medição experimental da massa dos íons em fase gasosa ^[36], a partir da razão entre a massa do íon e sua carga (m/z). Essa técnica é muito usada na caracterização de peptídeos, razão pela qual seu uso é largamente difundido na identificação de estruturas primárias de proteínas.

Um espectrômetro de massas é constituído basicamente por cinco componentes: a unidade de introdução de amostras, a fonte de ionização, o analisador de massas, o detector e o processador de sinais (**Figura 9**) ^[36].

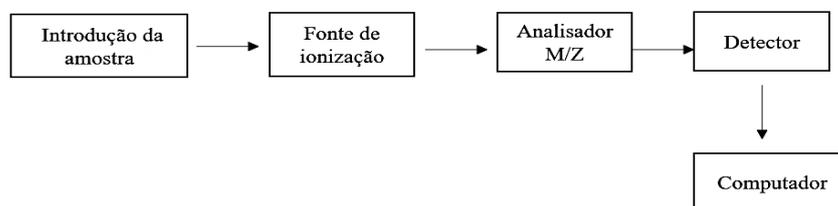


Figura 9 - Esquema simplificado de um espectrômetro de massa contendo seus principais componentes: fonte de íons, analisador m/z , detector, sistema de vácuo e o computador.

A análise de peptídeos é normalmente feita em modo positivo. Para que peptídeos, proteínas possam ser analisados usando espectrometria de massas é necessária uma forma branda de ionização para que a integridade da cadeia polipeptídica seja preservada. Os dois tipos de ionização mais usados na análise de proteínas são Ionização por Eletrospray ESI (sigla em inglês para ionização por eletrospray) e por MALDI (sigla em inglês para Ionização por Dessorção a Laser Assistida por Matriz) ^[36].

1.4.3 Alinhamento de sequências para o estudo de proteínas.

Devido à extensão de suas aplicações, o alinhamento de sequências biológicas é um processo de fundamental importância para a bioinformática e estudos relacionados a proteínas. Alinhamentos são técnicas de comparação entre duas ou mais sequências biológicas, que buscam séries de caracteres individuais que se encontram na mesma ordem nas sequências analisadas [24].

Durante o alinhamento, as sequências são organizadas em linhas e os caracteres biológicos integram as colunas do alinhamento. Seguindo à organização inicial, algoritmos específicos buscarão a melhor correspondência para as sequências em questão.

A ideia central destes algoritmos é minimizar as diferenças entre as sequências, buscando um alinhamento ótimo. Comumente, a similaridade entre as sequências envolvidas é expressa pelo termo identidade, que quantifica a porcentagem de caracteres idênticos entre duas ou mais sequências [24].

A relevância e abrangência do uso do método tornam os procedimentos de alinhamento o cerne para diferentes campos dentro da grande área da bioinformática.

Além de fundamentais em pesquisas de filogenética e análise evolutiva, os alinhamentos são úteis em estudos de inferência estrutural e funcional de proteínas, análises de similaridade e identificação de sequências e em estudos aplicados ao campo da genômica [24].

O BLAST sigla do inglês *Basic Local Alignment Search Tool* que significa ferramenta básica de pesquisa de alinhamento local constitui uma família de programas, que podem ser usados para alinhamento de sequências para diferentes fins, dependendo das necessidades do usuário. Esses programas variam quanto ao tipo de sequência de interesse, o banco de dados a ser empregado, e o tipo de comparação a ser realizada [24,56].

Esses programas são capazes de realizar buscas baseadas em alinhamento que, apesar de não serem exatas, são confiáveis e muito rápidas, sendo estas suas vantagens em relação a outros métodos [24,56].

1.4.4 Dicroísmo circular

O dicroísmo circular (CD), sigla do inglês circular dichroism, é uma técnica espectroscópica utilizada para estudar uma variedade de moléculas quirais, tais como fármacos, polímeros e biopolímeros, em solução. No caso das proteínas, o Dicroísmo Circular exerce importante papel na busca pelo conhecimento da estrutura e função biológica. Tais informações, por sua vez, são essenciais na busca por novos compostos com potencial terapêutico [24].

No CD, a energia ultravioleta é usada para efetuar análises e inferir informações estruturais dos compostos de interesse. O espectro de CD é gerado pela diferença na capacidade de absorção dos componentes esquerdo e direito da luz circularmente polarizada por moléculas quirais [24,39].

Os diferentes tipos de estrutura secundária de proteínas possuem espectros de CD bem característicos, estabelecidos a partir de modelos de oligo- e polipeptídios com estrutura secundária conhecida. Os espectros de CD de α hélices, folhas β e estruturas irregulares estão representados na (Figura 10) [24].

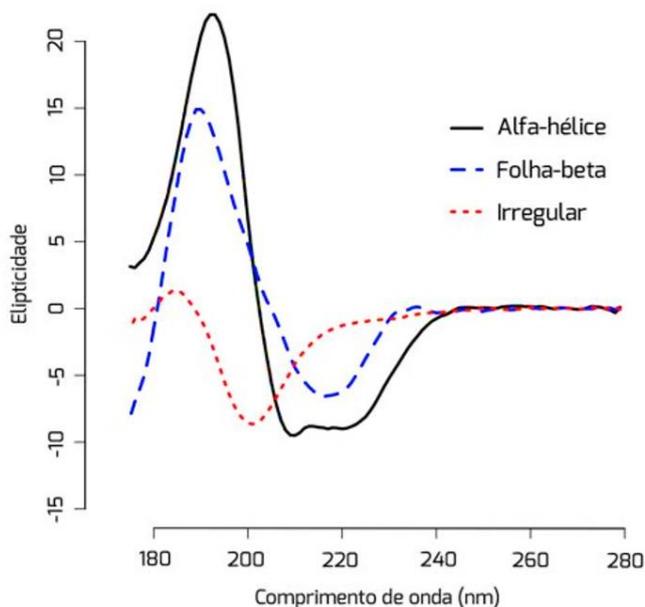


Figura 10- Espectros de CD típicos de estruturas do tipo α -hélices, folhas- β e estruturas irregulares [24].

Para fazer análises relativas às estruturas secundárias de proteínas, é necessário que o espectro adquirido no experimento de CD seja decomposto nos componentes hélice α , folhas β e estruturas irregulares e comparado a um banco de dados de proteínas com estruturas secundárias conhecidas, esse processo é conhecido como deconvolução espectral. A (Figura 11) a seguir apresenta o espectro de CD da albumina humana e a sua deconvolução [24].

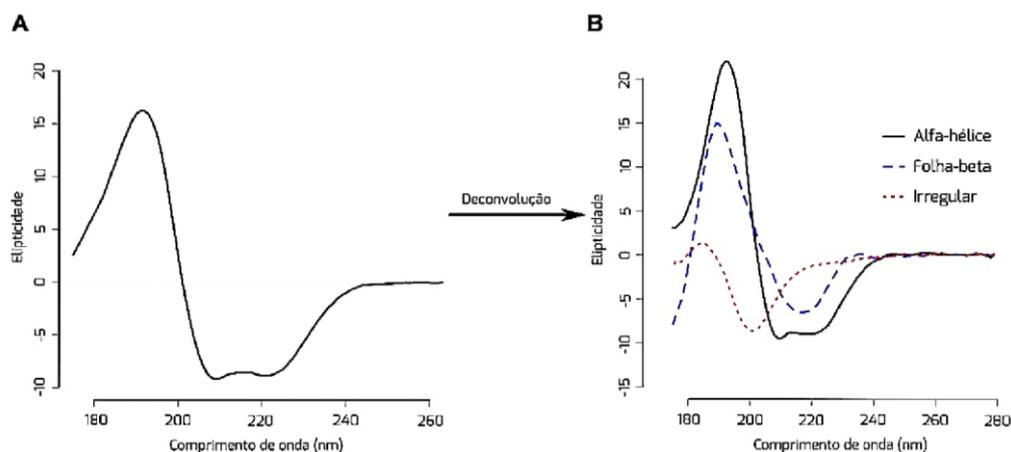


Figura 11 - Esquema da deconvolução espectral da albumina sérica humana [24]

1.4.5 Proteínas em plantas

As plantas possuem um arsenal de defesa composto por biomoléculas dentre elas proteínas e peptídeos relacionados à patogêneses (PR do inglês pathogenesis related), incluindo os peptídeos antimicrobianos (AMPs), no geral esses compostos não são tóxicos para mamíferos [37].

Essas proteínas e peptídeos são moléculas pequenas com massas abaixo de 10 kDa, que são classificadas em famílias de acordo com suas características bioquímicas e fisiológicas: Inibidores de protease, Defensinas, Tioninas e Proteínas transportadoras de lipídeos não-específicas (nsLTP) [37].

Os AMPs, são encontrados desde os organismos simples, como bactérias, até organismos mais complexos, como animais, incluindo-se os humanos, sendo assim considerados um mecanismo de defesa primitivos.

As PR's são moléculas definidas como proteínas que são induzidas em patologias ou situações estresse, e que atuam na defesa vegetal contra um confronto provável [38]. Atualmente, essa terminologia corresponde a um grupo de proteínas heterogêneo, codificadas por genes que são rapidamente induzidos por infecções patogênicas. Essas proteínas são solúveis em água, possuem pH's ácidos e são resistentes a proteólise [39].

Os peptídeos tais como tioninas (MM ~ 5 kDa), defensinas (MM ~ 5 kDa) e nsLTPs (MM 7 e 9 kDa), classificados como PRs, são peptídeos antimicrobianos (AMPs). AMPs são componentes importantes da defesa natural de diversos organismos

contra a invasão de patógenos. Esses peptídeos são anfipáticos, e apresentam comprimentos, sequências e estruturas variáveis, presença de um número variável de resíduos de cisteínas, que contribuem para estabilizar a conformação nativa pela formação de ligação de dissulfeto, além de espectro de ação contra uma ampla gama de [40,41,42].

Os AMPs já descritos têm sido divididos em vários grupos baseados em seu tamanho, estrutura secundária e terciária, e na presença ou ausência de pontes dissulfeto. A seguir algumas dessas AMPs serão descritas.

As Tioninas são pequenas proteínas (~ 5 kDa), ricas em resíduos contendo enxofre (cisteína, arginina e lisina), geralmente básicas ($pI > 8$), encontradas em monocotiledôneas e dicotiledôneas. Apesar de pequenas variações de comprimento (45-47 aminoácidos), eles compartilham a mesma estrutura tridimensional^[43].

A toxicidade das Tioninas a diferentes tipos de organismos e culturas de células tem sido investigada há várias décadas, após os relatos iniciais de suas propriedades antibióticas sugerindo uma importante função na defesa de plantas contra patógenos^[44].

As tioninas são sintetizadas após estresse e podem ser encontradas ou são expressas constitutivamente, podendo ser encontradas nas sementes (ou nos tecidos vegetativos, como no tonoplasto e na parede celular de folhas e flores, atuando como uma primeira linha de defesa^[45]).

Defensinas fazem parte do sistema imune inato do hospedeiro, uma estratégia de defesa antiga, usada por organismos multicelulares que incluem plantas, animais, e até mesmo fungos para controlar sua microbiota natural ou combater patógenos microbianos [46]

Defensinas vegetais possuem quatro pontes dissulfeto, apresentam entre 45 e 55 resíduos de aminoácidos, resultando em uma pequena molécula com massa entre 5 e 7 kDa^[47].

As Proteínas transportadoras de lipídeos (LTPs) são pequenas proteínas ricas em cisteína, encontrados em várias espécies de plantas. As LTPs são capazes de transferir vários tipos de lipídios, incluindo fosfatidilinositol, fosfatidilcolina e galactolipídeos, portanto apresentam baixa especificidade para o substrato lipídico, LTPs de plantas e devido a isso também são conhecidos como "proteínas transportadoras de lipídios não-específicas" (nsLTP)^[48,49].

Essas proteínas já receberam a classificação de acordo com sua massa molecular que consiste em duas subfamílias, LTP1 e LTP2, de acordo com a massa molecular, cerca de 9 e 7 kDa respectivamente [50,51]. Recentemente, a classificação dessas proteínas foi ampliada a nove diferentes classes por meio de uma análise filogenética.

Estruturalmente as nsLTPs possuem oito resíduos de cisteínas conservados, formando quatro pontes dissulfeto, que são responsáveis por sua estrutura e por sua estabilidade a temperatura e proteólise, propriedades que os tornam verdadeiros alérgenos alimentares [52].

A estrutura das nsLTPs formam uma cavidade hidrofóbica onde podem se alojar ácidos graxos, acil-CoA, ou fosfolípidios que poderiam ser transportados através da membrana celular das plantas [48]. Essas proteínas se encontram em um ambiente extracelular e por esta razão o transporte intracelular de lípidos é considerado improvável. Um peptídeo sinal, que os direciona as nsLTPs para a matriz extracelular faz parte da estrutura dessas proteínas [53].

As nsLTPs possuem estruturas bastante conservadas mas apesar disso podem desempenhar diferentes funções biológicas tais como envolvimento na mobilização de lípidos de armazenamento em sementes, formação da cutícula, na morte celular programada do endosperma, e na flexibilização da parede celular [54]. Essas proteínas apresentam capacidade de inibição de bactérias e fungos, pode ser considerada um inibidor de protease e sinalização da resistência sistêmica [55,56].

2 MATERIAIS E MÉTODOS

A metodologia empregada para isolamento e caracterização das proteínas de interesse dessa pesquisa foi desenvolvida em colaboração com o Instituto de Biologia da Universidade de Brasília nos laboratórios de Biofísica molecular, no laboratório de Toxicologia A (**Figura 12**) apresenta o fluxograma das etapas realizadas para isolamento e caracterização das proteínas de interesse. Cada etapa será descrita detalhadamente ao longo deste capítulo.

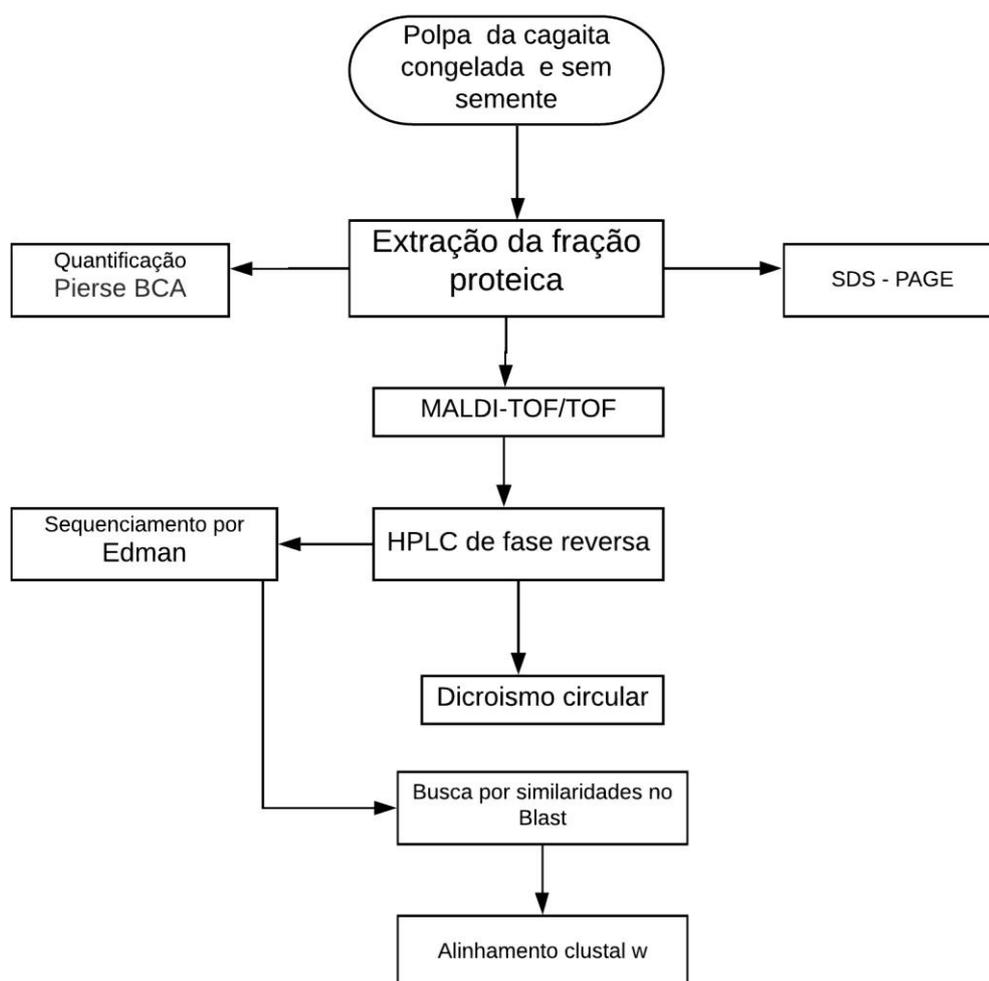


Figura 12 - Fluxograma da metodologia realizada para isolamento e caracterização das proteínas presentes na fração proteica do fruto da cagaiteira.

2.1 Extração da fração proteica

Foram utilizados frutos maduros congelados de *Eugenia dysenterica* coletados em setembro e outubro de 2017 nas cidades de Planaltina-DF e Brasília-DF. O protocolo de extração utilizado foi adaptado do trabalho de Lima et al. (2010) [22]. Esse protocolo consistiu em liquidificar as polpas sem casca e sementes e acondicioná-las a uma temperatura de aproximadamente -20 °C. Posteriormente, o material liquidificado foi suspenso em uma solução de NaCl (0,6 M) e HCl 0,1 % na proporção 1:3 (m/v). Em seguida, o extrato foi submetido à agitação por aproximadamente 16 horas a 4 °C. Este extrato foi centrifugado a 2.500 giros por 90 min e, em seguida, o sobrenadante foi precipitado em temperatura ambiente com sulfato de amônia até que se atingisse 100 % de saturação. O material foi deixado em repouso para que o excesso de

sulfato de amônio decantasse sendo, em seguida, centrifugado a 14.400 giros até a formação do *pellet*. O *pellet* foi resuspenso em água e dialisado por 8 vezes utilizando um Amicon ultra de 15 mL com *cutoff* de 3 kDa. Após esta etapa, o extrato bruto obtido foi acondicionado a uma temperatura de aproximadamente -20 °C para utilização posterior.

2.2 Quantificação da fração proteica

Para quantificação da fração proteica de *Eugenia dysenterica* foi utilizado o método Pierse BCA^[63]. Inicialmente pipetou-se 25 µL de amostra e 200 µL do reagente de Pierse BCA da marca Sigma-Aldrich em uma placa, as amostras ficaram incubadas por 30 minutos a 37°C, conforme instrução do fabricante. Em seguida, realizou-se a leitura da reação em comprimento de onda de 562 nm em um espectrofotômetro Jasco V-530 (Jasco Corporation). Para determinar a concentração proteica foi usada uma curva de calibração construída com soluções padrão de albumina sérica bovina (BSA, 0.1-1mg/mL; Sigma Steinhem, Germany). As leituras foram feitas em triplicatas a temperatura ambiente (25°C).

2.3 Análise da fração proteica

As proteínas constituintes da fração proteica da polpa madura da cagaita foram analisadas por meio de SDS-PAGE tris-tricina 15% e SDS-PAGE tris-glicina 15% descritas por Schagger e por Von Jagow (1987)^[56] e Laemmli (1970)^[57] respectivamente, com algumas modificações. Em ambas as metodologias, as amostras foram ressuspensas em tampão de amostra Laemmli 1x (200 mM Tris-HCl pH 6,8, SDS 10% (m/v), 10 mM β-mercaptoetanol, azul de bromofenol 0,05% (m/v) e 40% de glicerol). Depois de ressuspensas, as amostras foram fervidas a aproximadamente 95 °C em banho-maria por 15 min. No primeiro método, o gel separador de 15% (m/v) foi preparado utilizando-se 3 µL de TEMED, 150 µL de SDS 10% (m/v), 50 µL de PSA (persulfato de amônia) 10 % (m/v), 1,2 mL de Tris-HCl pH 8,8, 2 mL de acrilamida 30% (m/v) e 1,7 mL de água destilada. O gel concentrador 4 % foi preparado utilizando-se 3 µL de TEMED, 150 µL de SDS 10% (m/v), 50 µL de PSA (persulfato de amônia) 10% (m/v), 1,25 mL de Tris-HCl pH 8,8, 2 mL de acrilamida 30% (m/v) e 1,2 mL de água destilada. A amostra (20 µL) foi adicionada à canaleta do gel, o qual foi corrido em tampão de corrida de cátodo

(Tris-HCl 25 mM, tricina 200 mM e SDS 0,1% (m/v) pH 8,2) e de ânodo (Tris base, pH 8,9) a aproximadamente 160 V e 20 mA.

No segundo método, o gel separador 12% (m/v) foi preparado utilizando-se 3 µL de TEMED, 150 µL de SDS 10% (m/v), 50 µL de PSA (persulfato de amônia) 10% (m/v), 1,25 mL de Tris-HCl pH 8,8, 2 mL de acrilamida 30% (m/v) e 1,7 mL de água destilada. O gel concentrador 5% (m/v) foi preparado utilizando-se 75 µL de SDS 10% (m/v), 3 µL de TEMED, 25 µL de PSA 10% (m/v), 340 µL de Tris-HCl pH 6,8, 420 µL de acrilamida 30% (m/v) e 1,75 mL de água destilada. A amostra (20 µL) foi adicionada à canaleta do gel, o qual foi corrido em tampão de corrida (Tris-HCl 25 mM, glicina 200 mM e SDS 0,1% (m/v)) a aproximadamente 160 V e 45 mA. Após a eletroforese, foram usados dois métodos para corar os géis: coloração com solução de Commassie (Commassie brilliant blue R-250 0,1% (m/v), etanol 50% (v/v) e ácido acético 10% (v/v)) por 12 h e descorado com ácido acético 20% (v/v). O outro método usado para a visualização das bandas proteicas no gel de poliacrilamida foi a coloração por nitrato de prata [58] Foi utilizado como padrão o marcador molecular *SeeBlue™ Plus2 Pre-stained Protein Standard* com uma mistura de 10 proteínas com massa entre 3 e 198 kDa.

Para análise do perfil proteico também foi utilizada a técnica de tipo MALDI-TOF/TOF em equipamento modelo AB SCIEX TOF/TOF™ 5800 (AB Sciex, Framingham, MA, EUA). O procedimento foi realizado no modo refletido positivo, sendo a faixa m/z analisada de 4.000 a 20.000 Da. A matriz usada para a análise foi o ácido sinapínico.

2.4 Purificação da fração proteica

Após a avaliação da efetividade da extração proteica, alíquotas de 10 mg da fração proteica foram ressuspendidas em 1 mL de solução 0,1% (v/v) de *trifluoroacetic acid* (TFA) em água Milli Q e em seguida as amostras foram centrifugadas a 14.000 giros e o sobrenadante foi filtrado em filtro para seringa de 0,45 µm. Alíquotas de 200 µL do sobrenadante foram aplicadas em uma coluna de fase reversa C18 (SShimadzu VP ODS) com fluxo de 1 mL/min e detecção em 216 nm. A coluna foi previamente equilibrada com solvente A (TFA 0,1%, v/v, em água Milli Q). As frações foram eluídas aplicando-se as seguintes condições: gradiente de 0 a 60% de solvente B (TFA 0,1%, v/v, em acetonitrila) as corridas tiveram duração de 50 min. Após cada corrida a coluna foi lavada com 100 % de solvente B por 10 min. As frações de interesse

foram coletadas manualmente em tubos de 1,5 mL e secas em *speed vacuum* a 25°C e acondicionadas a -20°C. As frações obtidas nas purificações foram analisadas em espectrômetro de massas do tipo MALDI-TOF/TOF modelo AB SCIEX TOF/TOF™ 5800 (AB Sciex, Framingham, MA, EUA). O procedimento foi realizado no modo refletido positivo, sendo a faixa m/z analisada de 3.000 a 18.000 Da, a matriz usada para a análise foi ácido sinapínico. Depois de verificada a efetividade das purificações as amostras seguem para a etapa de caracterização biofísica e bioquímica. Nas purificações identificadas como parciais foi adicionada uma etapa de recromatografia dessas frações. Nas recromatografias as frações 1 e 2 foram eluídas aplicando-se as seguintes condições: gradiente de 0 a 30 % de solvente B (TFA 0,1%, v/v, em acetonitrila) nos 10 min iniciais das corridas, depois desse tempo foi mantida a porcentagem de 30 % (porcentagem de eluição ótima dos picos de interesse) as corridas tiveram duração de 25 min. A fração 3 foi eluída aplicando-se, gradiente de 0 a 35 % de solvente B (TFA 0,1%, v/v, em acetonitrila) nos 10 min iniciais das corridas, depois desse tempo foi mantida a porcentagem de 35% (porcentagem de eluição ótima dos picos de interesse) as corridas tiveram duração de 20 min. Por fim as frações 4 e 5 foram eluídas aplicando-se as seguintes condições: gradiente de 0 a 42% de solvente B (TFA 0,1%, v/v, em acetonitrila) nos 10 min iniciais das corridas, depois desse tempo foi mantida a porcentagem de 42% (porcentagem de eluição correspondente com os picos de interesse) as corridas apresentaram duração de 25 min. A efetividade das recromatografias foi verificada pelo mesmo método usado para as cromatografias iniciais.

2.5 Sequenciamento por Edman e análise da sequência obtida

A estrutura primária dos peptídeos foi determinada por degradação automática de Edman, utilizando-se o sequenciador Shimadzu PPSQ 33 A. Antes de cada análise, o sistema cromatográfico foi calibrado com PTH-aminoácidos padrões.

Foram realizadas buscas por similaridades à sequência primária parcial da proteína de 7,2 kDa utilizando o algoritmo blastp (NCBI) contra o banco de dados de sequências de proteínas não redundantes (nr) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>)^[59].

A sequência parcial obtida foi comparada à base de dados CDD (*Conserved Domain Database*) a fim de se observar a presença de domínios funcionais conservados. (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?RID=16NWYJ4201R&mode=all>)^[60].

O software clustal o foi utilizado para fazer o alinhamento múltiplo entre as 10 sequências que apresentaram maior percentual de identidade com a sequência parcial da proteína de interesse (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) [61].

2.6 Análise de estrutura secundária e termo estabilidade por dicroísmo circular

As análises de estrutura secundária das proteínas de 7 kDa (0,14 mg/mL), 9 kDa (0,27 mg/mL) e 13 kDa (0,19 mg/mL) presentes na polpa da cagaita foram realizados por dicroísmo circular utilizando o espectropolarímetro Jasco J-815 (Jasco Corporação, Tóquio, Japão) equipado com um sistema de controle de temperatura do tipo Peltier, conectado a uma bomba de água (Instrumentos Analíticos, Japão). Os espectros dicroicos foram coletados utilizando-se cubetas de quartzo de 0,05 cm, na faixa de comprimentos de onda distantes da ultravioleta (190-260 nm) com intervalos de 0,2 nm, com uma taxa de escaneamento de 100 nm/min, na temperatura de 25 °C. As amostras proteicas foram preparadas nos tampões acetato de sódio 5 mM, pH 4,0 e tampão Tris-HCl 5 mM, no pH 7,0 e pH 9,0. Para cada pH analisado, foram registradas dez leituras consecutivas, e do espectro dicroico resultante da média entre as leituras foi subtraído a contribuição dos tampões utilizados nos ensaios. As elipticidades (mdg) obtidas foram convertidas em elipticidade molar ($[\theta]$) ($\text{grau.cm}^2.\text{dmol}^{-1}$) baseado na massa calculada por resíduo de 112 Da segundo a equação (1) [61].

$$[\theta] = 112 \times \alpha / 10 \times C \times l \text{ (eq.1)}$$

Onde, α é o sinal dicroico em miligraus, C é a concentração da amostra (mg. mL^{-1}) e l é o trajeto óptico (cm) da cubeta utilizada.

Os ensaios de termoestabilidade das proteínas de 7 kDa e 9 kDa foram realizados nos pHs 4,0; 7,0 e 9,0; nos comprimentos de onda de 208 nm, e de 218 nm para a proteína de 13 kDa o comprimento de onda utilizado foi, correspondendo às regiões mais estruturadas das proteínas em estudo. As temperaturas dos ensaios variaram de 25 a 95 °C, com taxa de aquecimento de 0,2 °C/min. Durante todo o ensaio foram obtidos espectros dicroicos a cada 10 °C (25 a 95 °C) na região distante do ultravioleta (190-260 nm).

3 RESULTADOS E DISCUSSÕES

Este capítulo apresenta os principais resultados obtidos durante os ensaios experimentais e as discussões pertinentes acerca deles.

3.1 Extração e quantificação da fração proteica

Após o processo de extração descrito no item 2.1, as amostras foram quantificadas visando definir a quantidade aproximada de proteínas totais no extrato. A (Figura 13) apresenta a curva de calibração utilizada para o cálculo das concentrações aproximadas.

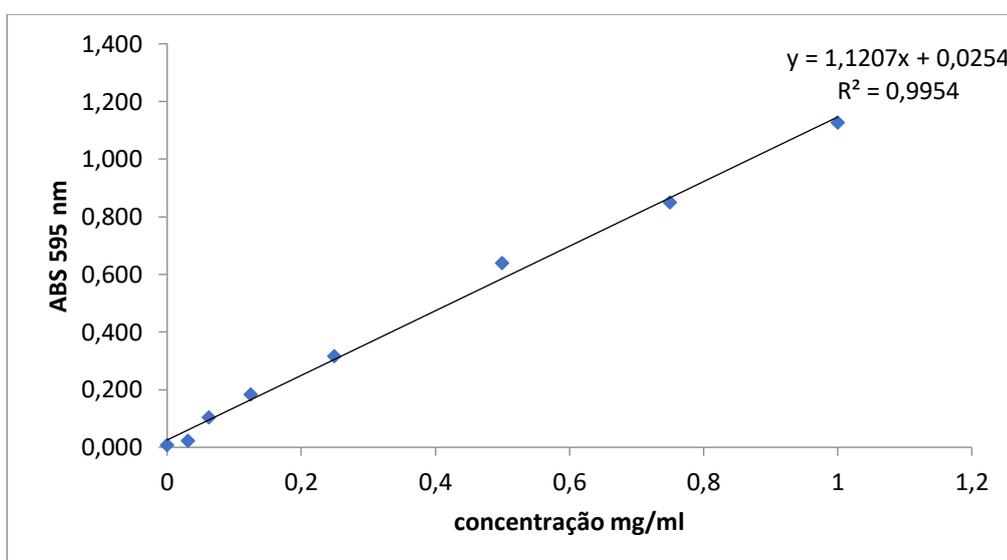


Figura 13 - Curva de calibração construída com soluções padrão de albumina sérica bovina para a quantificação de proteínas totais no material proveniente da extração da fração proteica do fruto da cagaiteira.

As extrações apresentaram uma concentração média de aproximadamente 0,4 mg/mL ou 24 mg /100g de fruto. O resultado obtido quando comparado com o descrito na literatura 30mg/100g de fruto ^[22], mostra que o rendimento médio das extrações foi de 80% as diferenças obtidas podem explicadas pelo uso de diferentes lotes de amostra que podem diferir dentre outras coisas no grau de maturação.

3.2 Análise da fração proteica por eletroforese em gel de poliacrilamida

A análise por eletroforese em gel de poliacrilamida foi feita inicialmente em gel desnaturante tris-tricina conforme descrito no item 2.3. Essa análise permitiu observar o perfil

das proteínas presentes no extrato bruto da fração proteica obtida no processo de extração. O gel de tris-tricina é apresentado na **(Figura 14)**.

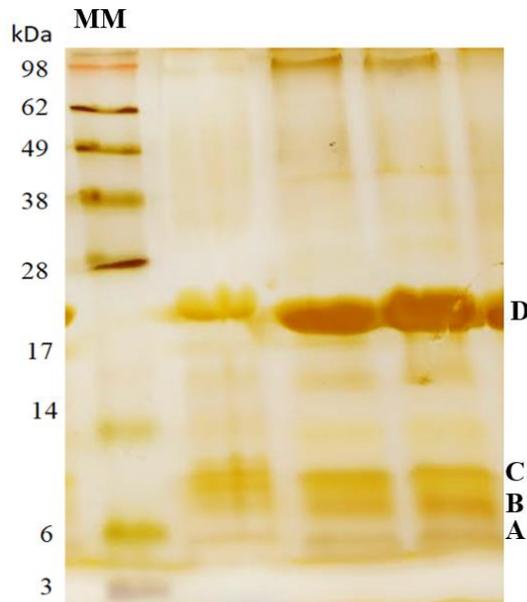


Figura 14 - Perfil da fração proteica em gel de tris-tricina corado com prata usando o marcador de massa molecular *SeeBlue™ Plus2 Pre-stained Protein Standard* com uma mistura de 10 proteínas com massa entre 3 e 198 kDa. Quatro bandas proteicas foram observadas, uma proteína de aproximadamente 7kDa (A), uma banda na altura de 9 kDa (B), uma proteína entre 13 e 14 kDa(C) e por fim pode ser observado uma banda de aproximadamente 20 kDa (D).

Um segundo método foi utilizado para análise do extrato bruto obtido da fração proteica obtida no processo de extração. O gel desnaturante de tris-glicina obtido neste ensaio está representado **(Figura 15)**.

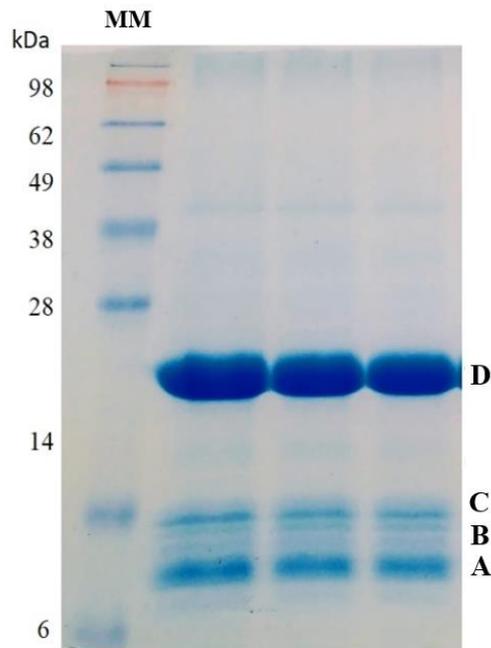


Figura 15 - Perfil da fração proteica em gel tris-glicina corado com comassie blue e usando o marcador de massa molecular *SeeBlue™ Plus2 Pre-stained Protein Standard* com uma mistura de 10 proteínas com massa entre 3 e 198 kDa.(MM) As 4 bandas correspondentes a uma proteína de aproximadamente 7 kDa (A), uma banda na altura de 9 kDa (B) estão evidenciadas. A seta amarela aponta uma proteína entre 13 e 14 kDa(C) e por fim pode ser observado uma banda de aproximadamente 20 kDa (D).

Em ambos os géis foram observadas bandas proteicas bem definidas. A primeira banda com aproximadamente 20 kDa, uma segunda de 9 kDa e uma 7 kDa, em ambas as figuras, também foi observada uma banda suave de aproximadamente 13 kDa. Os dois métodos foram satisfatórios para análise preliminar das proteínas presentes no extrato proteico da cagaita, portanto, optamos por trabalhar somente com o gel de tris-glicina, porque o processo de polimerização e as corridas feitas nesses géis foram mais rápidas.

O extrato proteico analisado por eletroforese também foi caracterizado por espectrometria de massas tipo MALDI-TOF/TOF conforme descrito no item 2.4 deste trabalho. O espectro representativo dos espectros obtidos é apresentado na (**Figura 16**).

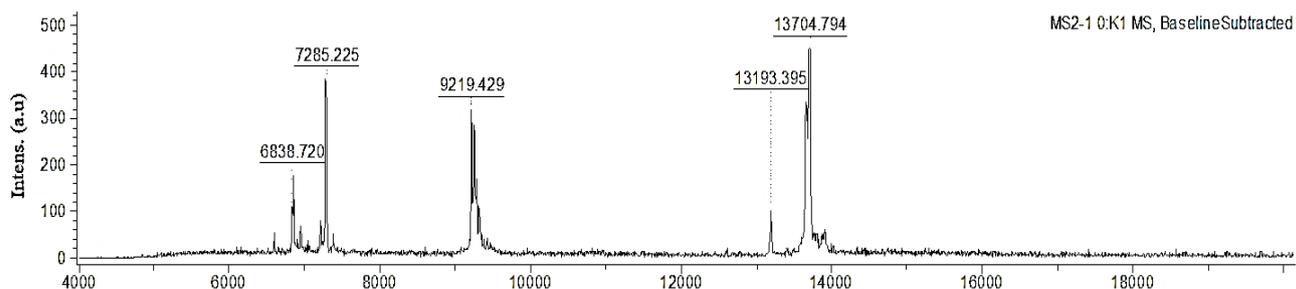


Figura 16 - Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico. Extrato aquoso do fruto de *Eugenia desynerica*, observa-se os componentes de massas aproximadas de 7,9 e 13kDa e o dupla carga do componente de massa aproximada de 13kDa.

Nos resultados obtidos nas análises do Maldi foi confirmada a presença das proteínas de 7,284 kDa, 9,252 kDa e 13,188 kDa, entretanto, a proteína de aproximadamente 22 kDa observada por eletroforese, não foi identificada nos espectros de massas, segundo estudos essa banda do gel pode ser proveniente da dimerização ou até mesmo da trimerização das outras proteínas presentes no extrato proteico ^[21].

3.3 Purificação da fração proteica

Alíquotas do extrato proteico foram aplicadas na coluna C18 (Shimadzu VP ODS) e detectadas em 216 nm. Foram divididas em 23 frações, que foram coletadas manualmente em tubos de 1,5 mL do tipo *ependorfs*. O perfil cromatográfico obtido nas corridas está mostrado na (**Figura 17**).

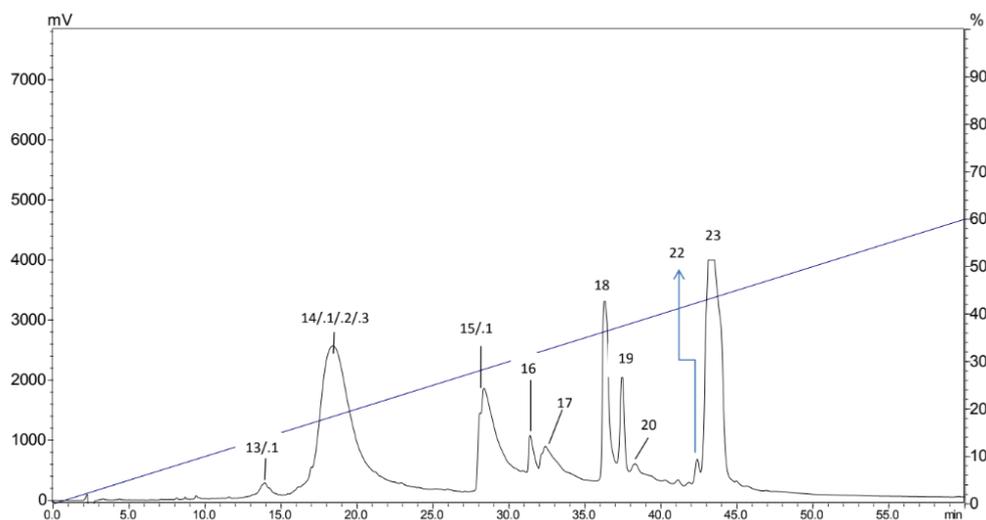


Figura 17- Corrida de triagem da fração proteica da cagaita foi usada uma coluna de fase reversa C18 (shimadzu VP ODS) e as amostras foram eluídas com gradiente de 0 a 60 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 50 minutos.

Após a corrida de triagem, as frações coletadas foram analisadas pela técnica de espectrometria de massas do tipo MALDI visando definir as frações que continham as proteínas de interesse. As frações 15.1, 16, 18, 19 e 23 coletadas no ensaio cromatográfico apresentavam as proteínas de interesse, conforme Tabela 1 as outras frações apresentavam misturas de vários componentes de massas por isso foram descartadas. Os espectros de massas obtidos nessa corrida de triagem encontra - se na sessão de anexos desse trabalho .

Tabela 1 - Frações contendo coletadas no ensaio de cromatografia HPLC que apresentava as proteínas de interesse e suas respectivas massas

Frações	Massa kDa
15.1 e 16	7,284
18	9,252
19 e 23	13,7

Nas corridas seguintes apenas as frações descritas na tabela acima foram coletadas. Nessas corridas foram coletadas manualmente uma média de 5 frações os perfis cromatográficos das corridas encontram-se na (**Figura 18**).

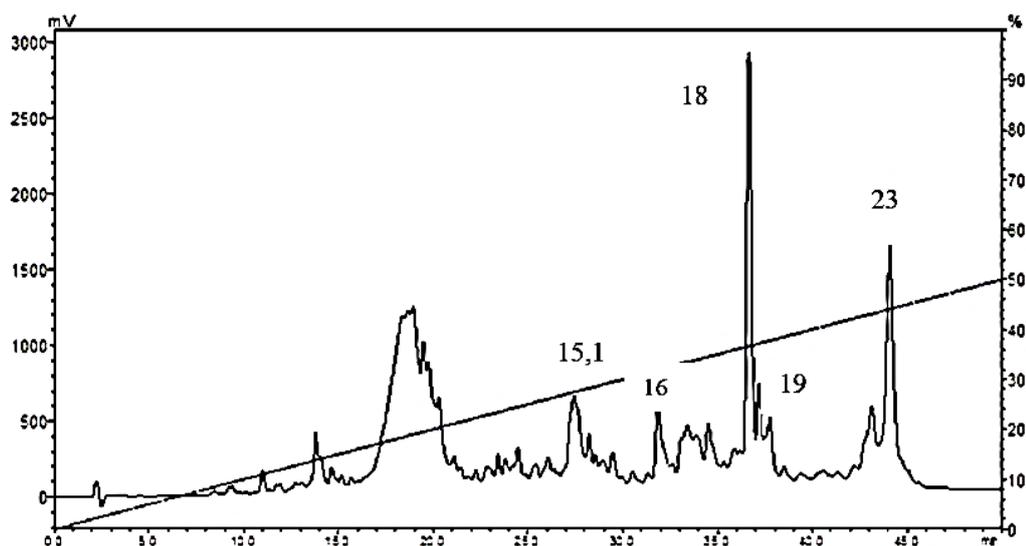


Figura 18 - Perfil cromatográfico das corridas feitas para isolamento das frações de interesse realizadas em uma coluna de fase reversa C18 (shimadzu VP ODS) eluídas com gradiente de 0 a 60 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 50 minutos.

Após o término das corridas de purificação, as amostras com semelhanças (frações com tempo de eluição semelhantes) seguiram para análise de massas.

Ao analisar os espectros das frações purificadas foi observado que houve uma purificação parcial das proteínas de interesse. No espectro coletado foi observada a presença de mais de uma das proteínas de interesse. Portanto, foi adicionada mais uma etapa de purificação visando obtenção de proteínas puras para os próximos ensaios. Essa etapa foi constituída por recromatografias das frações, a coluna usada nas cromatografias foi a mesma usada nas anteriores e novas condições cromatográficas foram testadas, com objetivo de isolar as frações de interesse. Os cromatogramas obtidos nas condições usadas na corrida encontram nas **(Figuras 19 a 21)**.

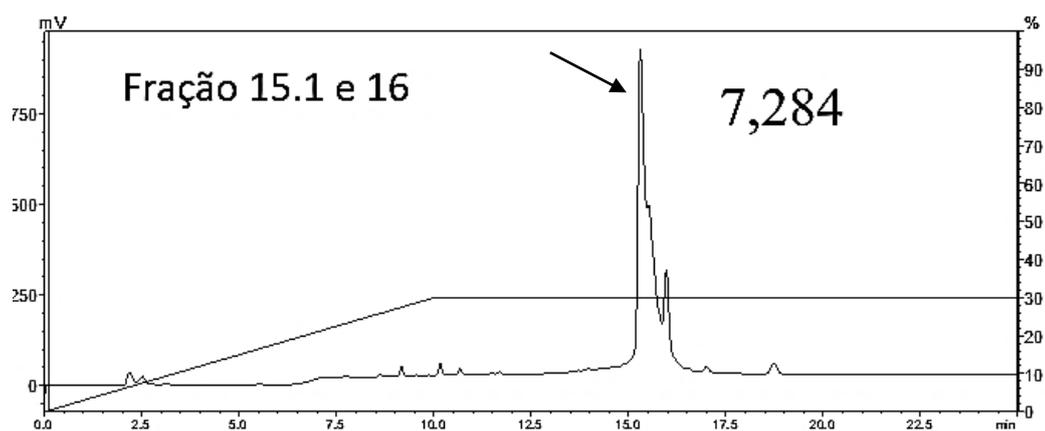


Figura 19 - Cromatograma representativo das recromatografias das frações 15.1 e 16 correspondente a proteína de 7,284 kDa as amostras foram eluidas em uma coluna de fase reversa C18 (shimadzu VP ODS) com gradiente de 0 a 30 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 25 minutos a fração de interesse coletada está destacada com uma seta.

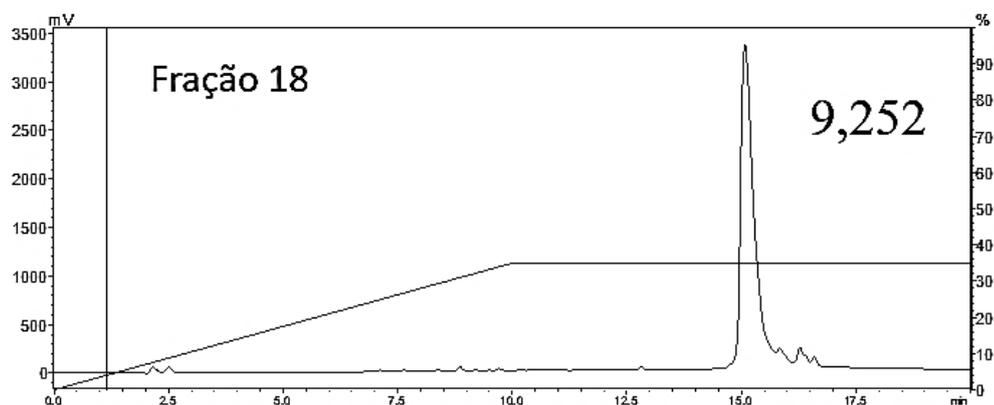


Figura 20 - Cromatograma representativo das recromatografias das frações 18 correspondentes a proteína de 9,252kDa, as amostras foram eluidas em uma coluna de fase reversa C18 (shimadzu VP ODS) com gradiente de 0 a 35 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 20 minutos a fração de interesse coletada está destacada com uma seta.

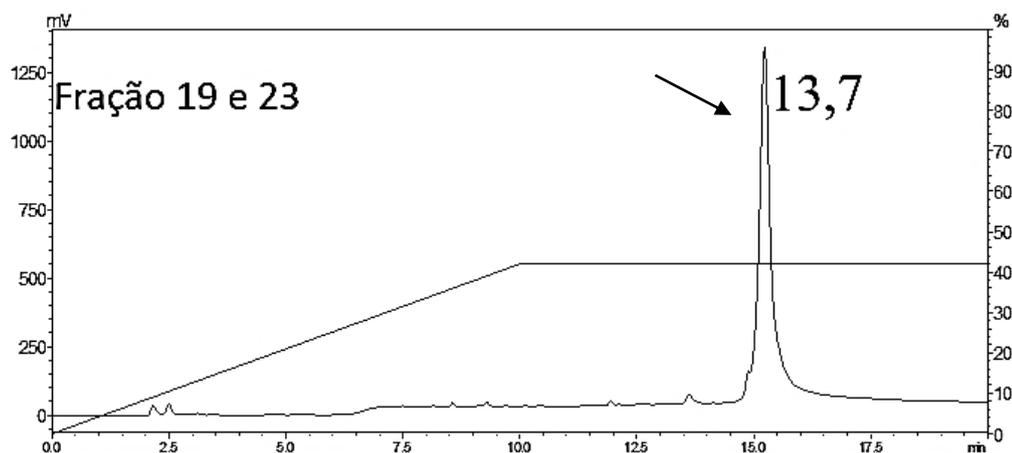


Figura 21 - Cromatograma representativo das recromatografias das frações 19 e 23 correspondente a proteína de 13,7 kDa as amostras foram eluídas em uma coluna de fase reversa C18 (shimadzu VP ODS) com gradiente de 0 a 42 % de solvente B (TFA 0,1 %, v/v, em acetonitrila) em corridas com duração de 20 minutos a fração de interesse coletada está destacada com uma seta.

Após as recromatografias e as análises de massas para comprovar a eficiência das purificações efetuadas (**Figuras 22 a 24**). Após a purificação as amostras foram secas e acondicionadas a -20 °C.

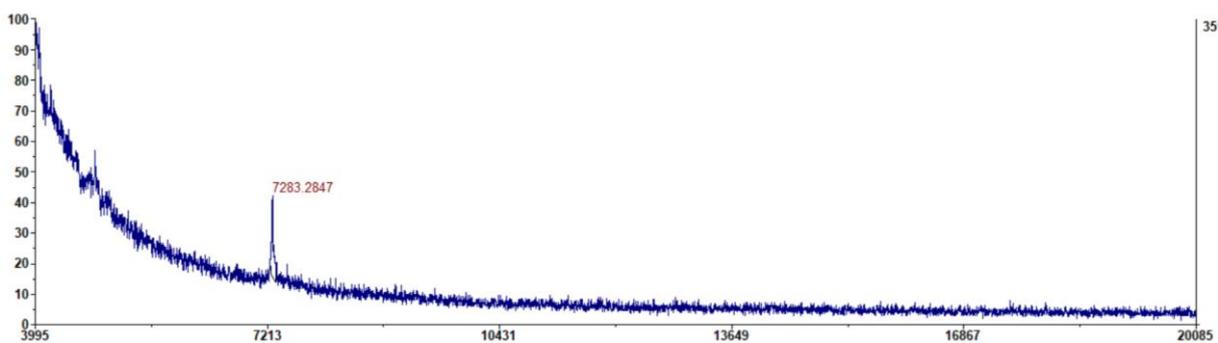


Figura 22- Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico, para a fração correspondente a proteína com massa aproximada de 7kDa.

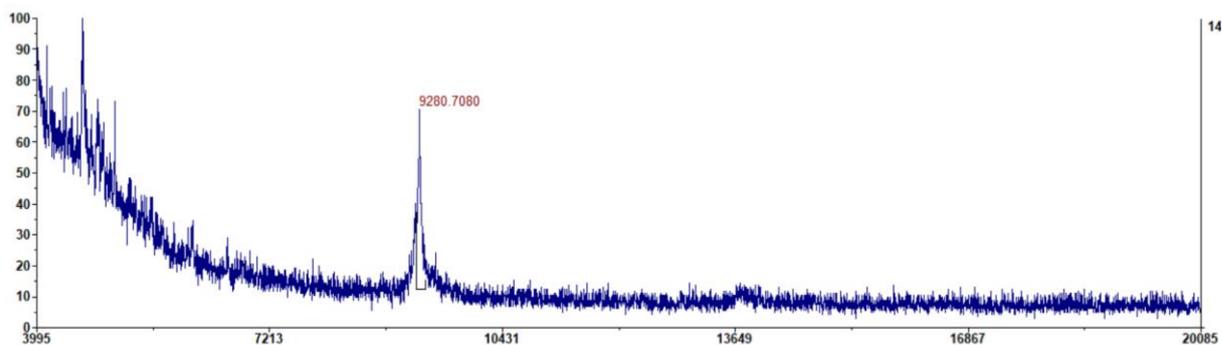


Figura 23 - Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico, para a fração correspondente a proteína com massa aproximada de 9kDa.

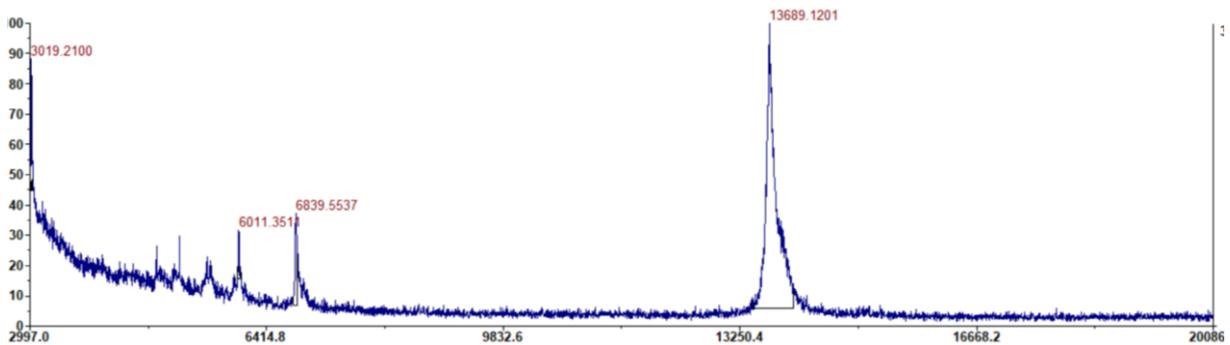


Figura 24 - Espectro de massa obtido em sistema de MALDI-TOF, modo linear positivo, matriz ácido sinapínico, para a fração correspondente a proteína com massa aproximada de 13kDa.

Na tabela a seguir é apresentado o resumo do rendimento das extrações e purificações.

Tabela 2 - Rendimentos extração e purificações

Quantidade de polpa	Amostra	Concentração da amostra (mg/mL)	Volume da amostra (mL)	Quantidades finais (mg)
590,5 g	Fração proteica	0,4	9,7	3,88
	Proteína 7 kDa	0,14	2	0,28
	Proteína 9 kDa	0,27	2	0,54
	Proteína 13 kDa	0,19	2	0,38

3.4 Sequenciamento por degradação de Edman

As amostras que foram purificadas seguiram para a caracterização bioquímica primária por degradação de Edman. Esta técnica foi usada para elucidar a estrutura primária das proteínas de interesse. Das amostras analisadas, somente a amostra com aproximadamente 7 kDa pode ser parcialmente sequenciada, as outras duas proteínas de interesse apresentaram a porção N-

terminal bloqueada. Para a proteína de 7,284 kDa foram sequenciados 38 resíduos de aminoácidos que estão representados pelo código de uma letra na (Figura 25).

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
V	T	X	S	A	T	E	L	S	S	X	V	S	A	I	T	S	S	A	P	P	S	A	L	X	X	S	K	L	R	E	Q	R	P	X	L	X	G

Figura 25 - Sequência parcial da proteína de 7,2 kDa obtida por degradação automática de Edman onde X são prováveis cisteínas.

O resultado da busca por similaridade e o alinhamento múltiplo mostrou identidade com 61 proteínas não específicas de transferência de lipídeos (nsLTP), sendo a maior identidade encontrada de 84,21% para *Syzygium oleosum* e *Eucalyptus grandis*. A proteína estudada possui uma massa molecular de 7,2 kDa correspondendo a aproximadamente 65 aminoácidos baseando-se na massa molecular média (Daltons) por aminoácidos e foi classificada como sendo uma nsLTP putativa de acordo com a similaridade com outros membros desta família (Figura 26).

nsLTP_P_granatunXP031373702.1	-----MKAAYIVVFLVLLAAEQAEIAEGVTCKPTELS	SCVSAITSSSTPP	45
nsLTP_G_tomentosumbTYH98342.1	-----MKKASYTAALSAAAVVVLLLAELAKVSMATCSP	QLS SCVSAITSSSPP	50
nsLT_P_G_maxNP001236829.1	-----MKTSY-ATLCIMLTLLLLGQSELLCMAAVTC	NALQLSACASAITSSSTPP	48
nsLTP_G_soja_XP028221195.1	-----MKTSY-ATLCIMLTLLLLGQSELLCMAAVTC	NALQLSACASAITSSSTPP	48
nsLTP_Q_suberXP023897419.1	-----MKKASY-VVLCVAVVAMLLFEAPLMAKAVTC	SPLQLS SCIAAITSSAPP	49
nsLTP_V_viniferaXP002284213.1	-----MKIPY-AALC--VIM-VVLLSEHLTKAVTC	SPLQLS SCFAAITSSAPP	45
nsLTP_A_chinensisPSR87834.1	-----MV-L--LVDVRVTEAVTC	SPTELS PCIGAITSSAPP	33
E. dysenterica	-----VTXSATELS	SXVSAITSSAPP	21
nsLTP_R_argenteaXP030527469.1	MTKPPSIRSPSPRLLA-AALVATALV-LLAPDRAPVAE	AVTCSPTELS SCASAITSSAPP	58
nsLTP_S_oleosumXP030444409.1	-MAKPSNRSPSPRLLV-AALVG-VLV-LLSVDQARVAE	AVTCATELS SCVSAITSSAPP	56
nsLTP_E_grandisXP010062205.1	-MAKPSDPSPSPRLLA-AALVA-VLV-LLSPDQARVAE	AVTCATELS SCVSAITSSAPP	56
		** . : ** . ***** **	
nsLTP_P_granatunXP031373702.1	SKLCCSKIREQKFC	LCQYLKPNPKLKFVNTFNARKVAQTCGTPPFRRC	92
nsLTP_G_tomentosumbTYH98342.1	SKLCCSKIKEQKFC	LCQYLKPNLKKFINTFNARKVASTCGTPPFC	97
nsLT_P_G_maxNP001236829.1	SALCCSKLKEQRF	CLCQYLKDPNLKLVNSFNARMAANACGSPFPTC	95
nsLTP_G_soja_XP028221195.1	SALCCSKLKEQRF	CLCQYLKDPNLKLVNSFNARMVANACGSPFPTC	95
nsLTP_Q_suberXP023897419.1	SSTCCTKLREQRF	CLCGYLKDPNLRQYVNSPGARKVASTCGVPPSC	96
nsLTP_V_viniferaXP002284213.1	SSMCCSKLREQRF	CLCGYIRDPNLSQYVNSANARRVASTCGVPPNC	92
nsLTP_A_chinensisPSR87834.1	SGTCCSKLREQRF	CLCGYLKPNLSQYVNSFNARRVATACGVPTPTC	80
E. dysenterica	SALXXSKLREQRF	XLXG-----	38
nsLTP_R_argenteaXP030527469.1	SALCCSKLREQRF	CLCGYIKPNLNRQYVTSFNARKRVARTCGVPSPTC	105
nsLTP_S_oleosumXP030444409.1	SALCCSKLREQRF	CLCGYIRNPNLRQYVTSFNARKRVARTCGVPPPTC	103
nsLTP_E_grandisXP010062205.1	SALCCSKLREQRF	CLCGYIRNPNLRQYVTSFNARKRVARTCGVPPPTC	103
		* : * : * : *	

Figura 26 - Alinhamento múltiplo da sequência primária parcial da proteína de E. dysenterica de 7kDa com 10 nsLTP's homólogas disponíveis no banco de sequências de proteínas não redundantes (nr). As proteínas estão identificadas com o número de acesso. Os resíduos conservados de Cys estão marcados de azul. Os demais resíduos conservados estão marcados de verde e os resíduos semi-conservados estão marcados de cinza.

A análise de domínios conservados da sequência obtida após comparação com a base de dados CDD (*Conserved Domain Database*) mostrou que a parte da proteína estudada possui um único domínio conservado localizado entre os resíduos Ser4 e Gly38, sendo este pertencente a subfamília das proteínas não específicas de transferência de lipídeos do tipo 2 (nsLTP2) como pode ser observado na (**figura 27**) com *E-value* $1,59 \times 10^{-6}$.

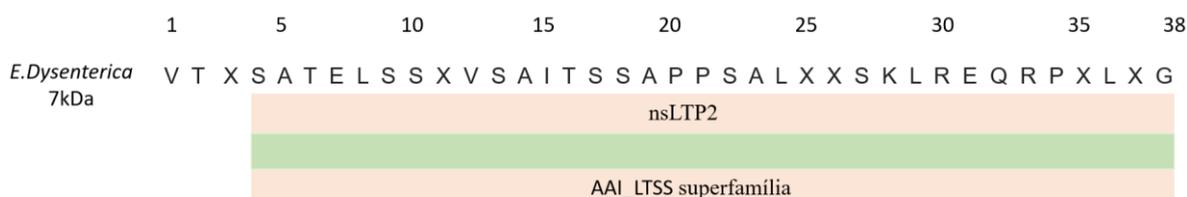


Figura 27 - Representação dos domínios conservados da sequência obtida após comparação com a base de dados CDD (*Conserved Domain Database*), que demonstra um único domínio conservado localizado entre os resíduos Ser4 e Gly38, sendo este pertencente a subfamília das proteínas não específicas de transferência de lipídeos do tipo 2 (nsLTP2)

Pode ser observado no alinhamento (**Figura 25**), que as Cys nas posições 38, 46, 60, 61, 70 e 71 (numeração da sequência de *Syzygium oleosum* e *Eucalyptus grandis*) são conservadas em todas as nsLTP analisadas; assim sendo, pode-se assumir que os resíduos de aminoácidos representados por X na sequência da nsLTP putativa de *Eugenia dysenterica* são de fato resíduos de Cys.

Ao calcular a massa da sequência parcial da proteína substituindo o x por Cys chegamos a massa aproximada de 3,858kDa esse valor se comparado com a massa total da proteína de interesse 7,285 kDa podemos concluir que aproximadamente 53% da cadeia polipeptídica foi sequenciada.

As nsLTPs (proteínas de transferência de lipídios não específicas) são proteínas pequenas e relativamente simples, com o tamanho variando de 6,5 a 10,5 kDa tal como a proteína parcialmente sequenciada nesse trabalho, são amplamente distribuídos no reino vegetal e estão presentes em abundância. Essas proteínas são sintetizadas como precursores com um peptídeo sinal secretório N-terminal, geralmente variando de 21 a 27 aminoácidos em tamanho. As nsLTPs apresentam estruturas estabilizadas por meio de ligações dissulfeto, o que torna essas proteínas resistentes ao calor e à proteólise^[49,50].

As nsLTPs são diretamente relacionados com casos de alergia alimentar, isso pode ser explicado pela alta estabilidade térmica e resistência à proteólise gastrointestinal dessas proteínas ^[51]. Muitos desses alérgenos das nsLTP's além de causar manifestação de reações alérgicas, também podem causar a sensibilização do trato gastrointestinal ^[51]

As características descritas das nsLTPs reforçam a hipótese da proteína de 7 kDa parcialmente sequenciada pode ser da família das nsLTPs, pois popularmente a polpa dos frutos de onde essa proteína foi extraída são conhecidos por sua capacidade laxativa, e esse efeito já foi comprovado cientificamente em testes em ratos ^[22].

3.5 Análise de estrutura secundária e termo estabilidade por dicroísmo circular

Proteína de 7 kDa presente na polpa de *E. dysenterica*

Os resultados obtidos nos ensaios de dicroísmo circular para a proteína de 7 kDa presente na polpa de *E. dysenterica* mostram que esta proteína possui estrutura secundária dependente de pH, apresentando espectros dicroicos característicos de proteína estruturada para todos os pHs testados, com sinal dicroico mais pronunciado no pH ácido (**Figura 28**).

Os espectros da proteína apresentam banda dicroica negativa principalmente em 208 nm e 222 nm correspondendo à estrutura de alfa hélice e uma banda positiva em 190 nm, seguido de uma banda dicroica negativa em 218 nm, região típica de estrutura de folha beta, e banda positiva em 195 nm (**Figura 28**).

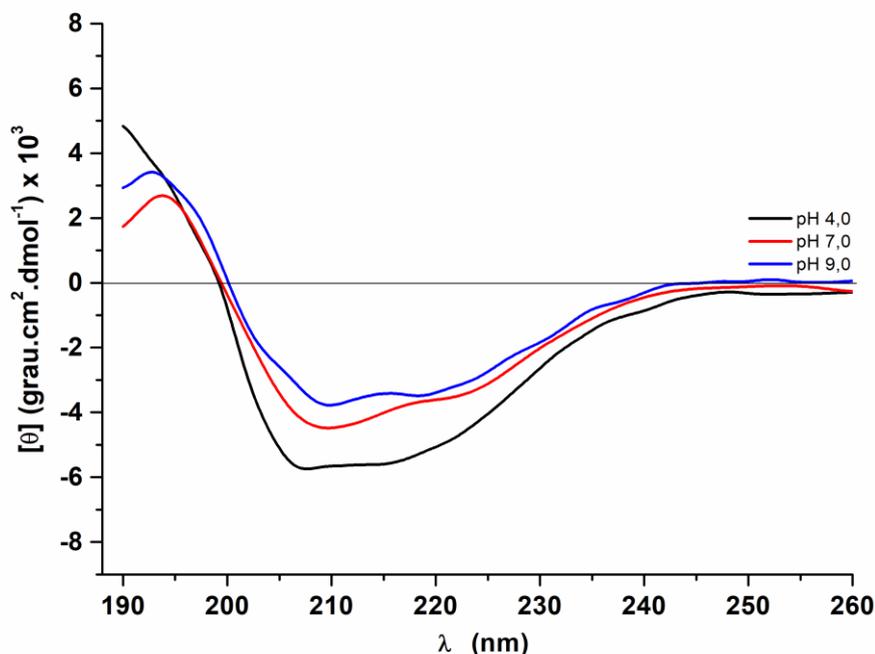


Figura 28 - Espectros dicróicos da proteína de 7 kDa na região distante da luz UV (190-260 nm) nos pH 4,0; 7,0 e 9,0 obtidos a 25 °C. A concentração proteica foi 0,19 mg/mL em solução nos tampões acetato de sódio 5 mM para o pH 4,0 e Tris HCl 5 mM para os pHs 7,0 e 9,0.

Estes resultados foram observados em todos os pHs testados, entretanto, no pH 4,0, o percentual de estrutura secundária é maior para alfa hélice (12%), em comparação aos pHs 7,0 e 9,0 (**Tabela 2**). Que apresentam redução no conteúdo de alfa hélice em ~2%, quando comparado ao pH ácido. Nos pHs 7,0 e 9,0, o conteúdo de folha beta paralela aumentou em 4%. Esses resultados mostram que a proteína de 7 kDa apresenta-se mais estruturada na faixa ácida isso pode ser explicado pelo fato de que o pH do cagaita madura é por volta de 3,0 ou seja é um fruto ácido^[38].

Tabela 3 - Conteúdo de estrutura secundária da proteína de 7 kDa na polpa de *E. dysenterica* em diferentes pHs obtidos por deconvolução dos espectros de DC utilizando o programa CDNN

Estrutura Secundária %	pH 4,0	pH 7,0	pH 9,0
α -hélice	12,1	10,8	10,5
β -antiparalela	28,5	32,3	33,7
β -paralela	5,4	5,4	5,5
Volta- β	20,0	19,4	19,1
Desordenada	34,4	33,9	33,5

A termo estabilidade da proteína de 7 kDa foi estudada a partir da obtenção de curvas de desnaturação com temperaturas variando entre 25 e 95 °C nos pHs 4,0, 7,0 e 9,0 (**Figura 28A**).

As curvas de desnaturação obtidas mostram que nos três pHs estudados houve somente manutenção no sinal dicróico com pequenas variações de intensidade em 208 nm variando de -5.600 para -4.300 grau.cm²/dmol no pH 4,0 e no pH 7,0 de -4.300 para -3.200 grau.cm²/dmol. No entanto, no pH 9,0 o sinal manteve-se próximo a -3.600 grau.cm²/dmol. Portanto, esse resultado mostra que as proteínas não foram desnaturadas nas condições testadas. Os espectros dicróicos coletados concomitantemente em diferentes temperaturas, mostram que mesmo após a elevação da temperatura a 95 °C, os espectros dicróicos apresentam as bandas negativas nas regiões de 208, 218 e 222 nm, o que caracteriza a proteína de 7 kDa de *E. dysenterica* como termoestável nos pHs estudados (**Figura 29B-D**). O fato da proteína ser termoestável é mais um indicio que de fato se trata se uma nsLTP que tem a estrutura estabilizada por 4 pontes de sulfeto ^[49].

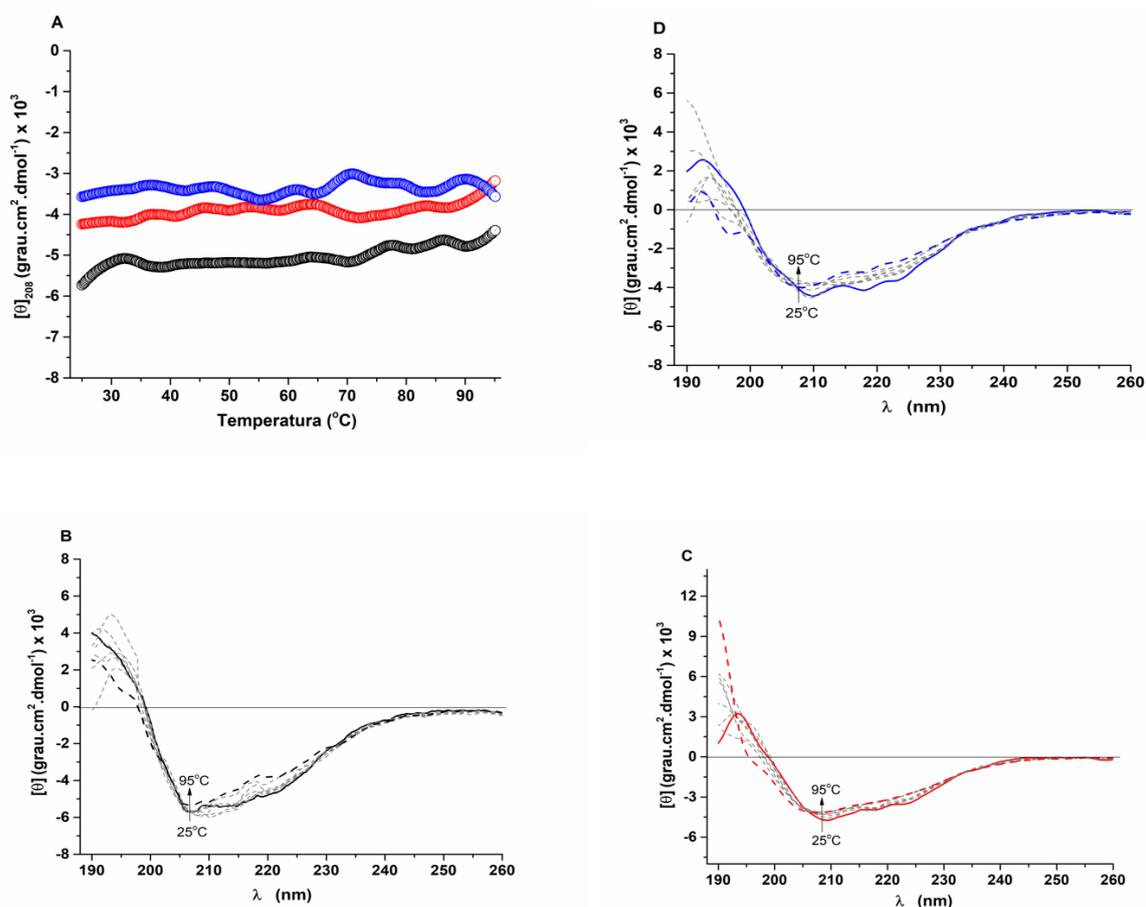


Figura 29 - Curvas de termoestabilidade e espectros dicróicos da proteína de 7 kDa. (A) e espectros dicróicos da proteína de 7 kDa presente na polpa de *E. dysenterica* nos pHs 4,0 (B), 7,0 (C) e 9,0 (D). Curvas de desnaturação térmica foram obtidas no comprimento de onda fixado em 208 nm (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul), com variação da temperatura de 25 °C á 95 °C. Os espectros dicróicos da proteína foram obtidos na região distante da luz-UV (198-260 nm), de 25 °C a 95 °C, com intervalos de 10 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul).

Proteína de 9 kDa presente na polpa de *E. dysenterica*

Os resultados obtidos nos ensaios de dicróismo circular para a proteína de 9 kDa presente na polpa de *E. dysenterica* mostram que a estrutura secundária é dependente de pH, apresentando espectros dicróicos característicos de proteína estruturada para todos os pHs testados, com sinal mais pronunciado nos pHs 7,0 e 9,0 (Figura 28). As bandas dicróicas negativas, principalmente em 208 nm e 222 nm, correspondem a estrutura de alfa-hélices, e a banda dicróica negativa em 218 nm e banda positiva em 198 nm correspondem à região típica de estrutura de folha beta, (Figura 30).

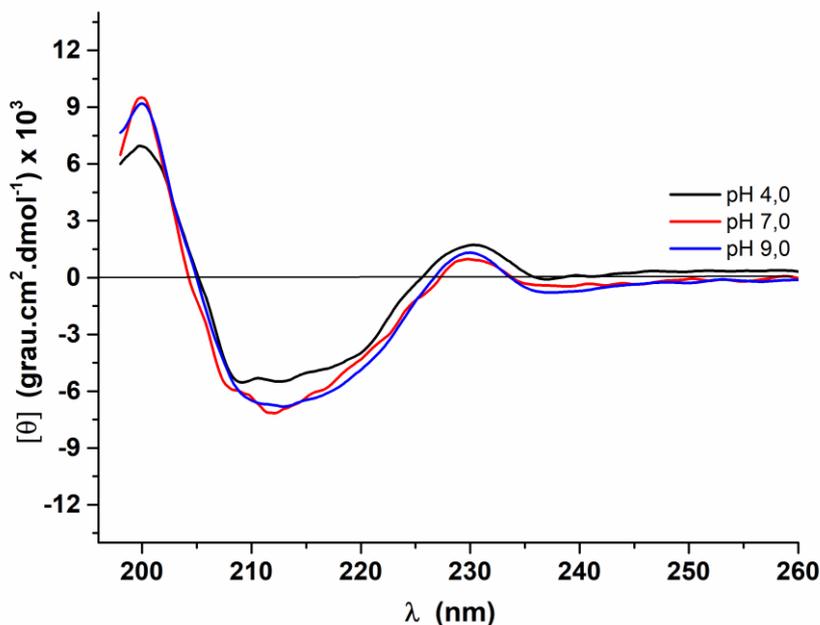


Figura 30 - Espectros dicróicos da proteína de 9 kDa na região distante da luz UV (198-260 nm) nos pHs 4,0; 7,0 e 9,0 com a concentração proteica de 0,27 mg/mL em solução nos tampões acetato de sódio 5 mM para o pH 4,0 e Tris HCl 5 mM para os pHs 7,0 e 9,0. Este resultado foi observado em todos os pHs testados, porém, nos pHs 7,0 e 9,0, o percentual de estrutura secundária é maior para alfa hélice (13,5%), em comparação ao pH 4,0 que é de 11,6% (Tabela 3). Nos pHs 7,0 e 9,0, o conteúdo de folha beta paralela reduz em ~2,5%. Esses resultados mostram que a proteína de 9 kDa apresenta-se mais estruturada na faixa neutra e básica.

Tabela 4 - Conteúdo de estrutura secundária da proteína de 9 kDa presente na polpa de *E. dysenterica* obtido por deconvolução dos espectros de DC por meio do programa CDNN.

Estrutura	pH 4,0	pH 7,0	pH 9,0
Secundária (%)			
α-hélice	11,6	13,9	13,5
β-antiparalela	35,9	32,6	33,6
β-paralela	6,4	6,6	6,6
Volta-β	15,3	14,5	14,3
Desordenada	33,7	32,6	32,6

A termo estabilidade da proteína de 9 kDa foi analisada a partir das curvas de desnaturação com temperaturas variando entre 25 e 95 °C nos pHs 4,0, 7,0 e 9,0 (**Figura 31A**). Estas curvas mostram que nos três pHs estudados houve redução no sinal dicróico em 208 nm variando de -5.000 para -2.500 grau.cm²/dmol no pH 4,0; no pH 7,0 de -6.000 para -2.000 grau.cm²/dmol, já no pH 9,0 o sinal variou de -5.700 para -3.000 grau.cm²/dmol. Esse resultado caracteriza um processo de desnaturação proteica parcial. Os espectros dicróicos coletados concomitantemente em diferentes temperaturas, mostram que até 85 °C houve uma redução gradual do sinal dicróico em 208 nm, 218 nm e 222 nm para os três pHs analisados. No entanto, após 85 °C, a banda dicróica negativa deslocou-se abruptamente para região de 200 nm nos pHs 4,0 e 9,0, caracterizando um aumento no teor de estrutura desordenada nessas faixas (**Figura 31-D**).

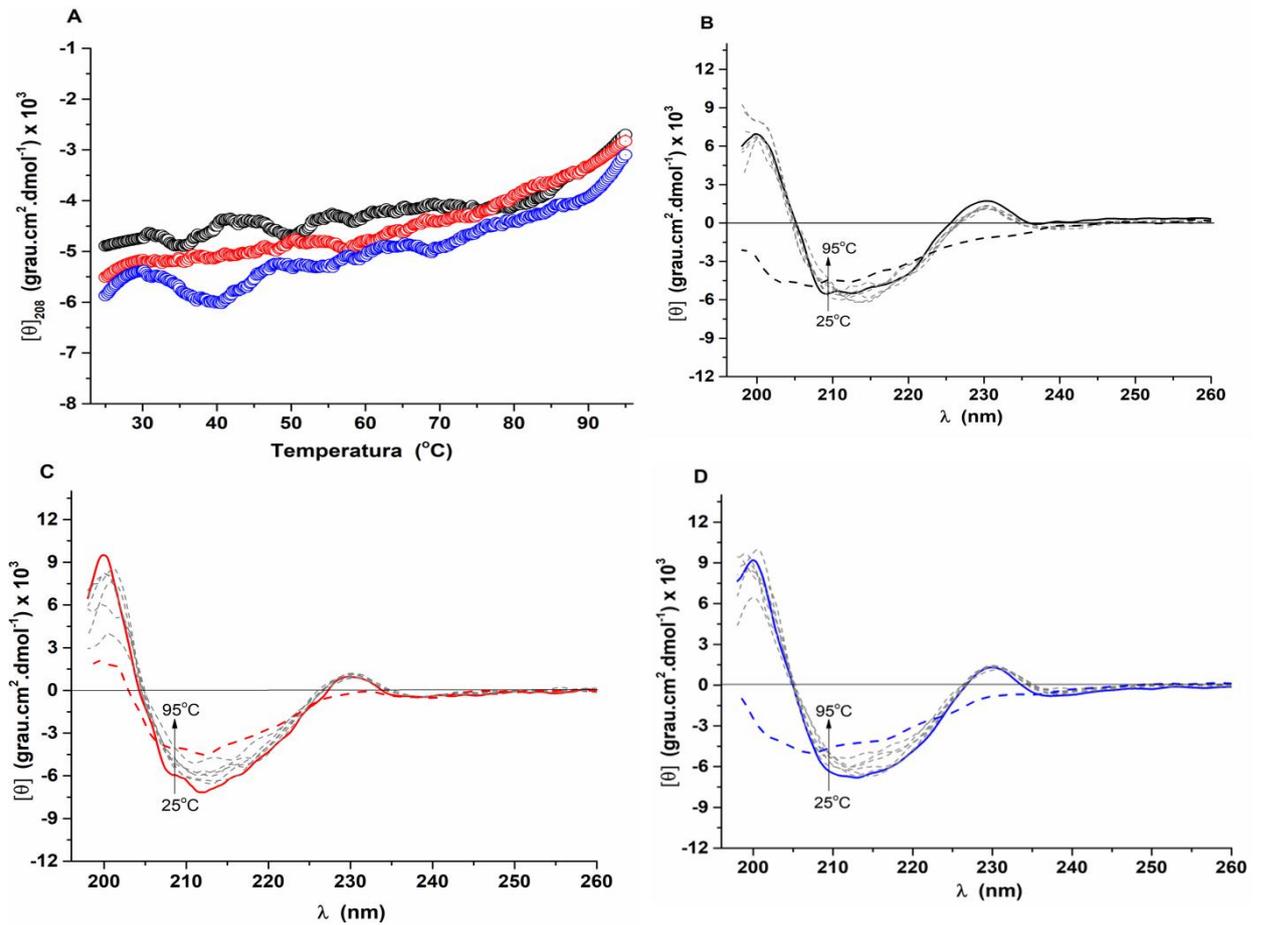


Figura 31 - Curvas de termo estabilidade e espectros dicróicos da proteína de 9 kDa. Curvas de termo estabilidade (A) e espectros dicróicos da proteína de 9 kDa presente na polpa de *E. dysenterica* nos pHs 4,0 (B), 7,0 (C) e 9,0 (D). As curvas de desnaturação térmica foram obtidas no comprimento de onda fixado em 208 nm, com variação da temperatura de 25 °C para 95 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul). Os espectros dicróicos da proteína foram obtidos na região distante da luz-UV (198-260 nm), de 25 °C a 95 °C, com intervalos de 10 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul).

Esse resultado mostra que a proteína de 9 kDa é mais termoestável em pH 7,0. A proteína em estudo desenvolveu-se parcialmente nos pHs testados, entretanto, não foi observado agregação na luz visível, portanto, um espectro dicróico foi coletado após ao ensaio de desnaturação ao retornar para 25 °C, a proteína aparentemente se reenovelou nos três pHs testados, apresentando um espectro dicróico bem próximo ao espectro inicial obtido a 25 °C (Figura 32).

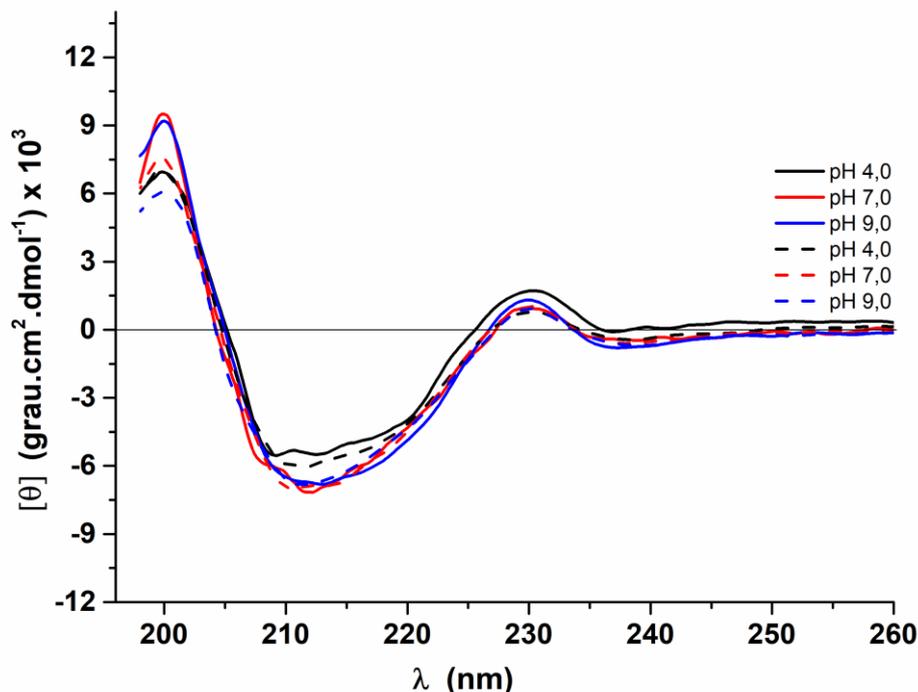


Figura 32 - Espectros dicróicos de desnaturação e de renaturação da proteína de 9 kDa *E. dysenterica* na região distante da luz UV nos pHs 4,0, pH 7,0 e pH 9,0 obtidos a 25 °C (linhas contínuas) e após ensaio de desnaturação térmica (25 a 95 °C) com retorno a 25 °C (linhas pontilhadas). Os espectros dicróicos da proteína de 9 kDa foram obtidos de 198-260 nm antes e após desnaturação térmica.

Proteína de 13 kDa presente na polpa de *E. dysenterica*

Os resultados obtidos nos ensaios de dicróismo circular para a proteína de 13 kDa presente na polpa de *E. dysenterica* mostram que a estrutura secundária é dependente de pH, apresentando espectros dicróicos característicos de proteína estruturada para todos os pHs testados, com sinal mais pronunciado nos pHs 4,0 e 7,0. Os espectros da proteína apresentam banda dicróica negativa principalmente em 208 nm e 222 nm correspondendo a estrutura de alfa-hélices, seguido de uma banda dicróica negativa em 218 nm, região típica de estrutura de folha beta, e banda positiva em 195 nm (**Figura 33**). O mesmo resultado foi observado em todos os pHs testados, contudo, nos pHs 4,0 e 7,0, o percentual de estrutura secundária é maior para alfa hélice (~12%), em comparação ao pH 9,0 que é de ~10% (**Tabela 5**). No pH 9,0, o conteúdo de folha beta paralela aumentou em 3,0%. Esses resultados mostram que a proteína de 13 kDa apresenta-se mais estruturada na faixa ácida e neutra.

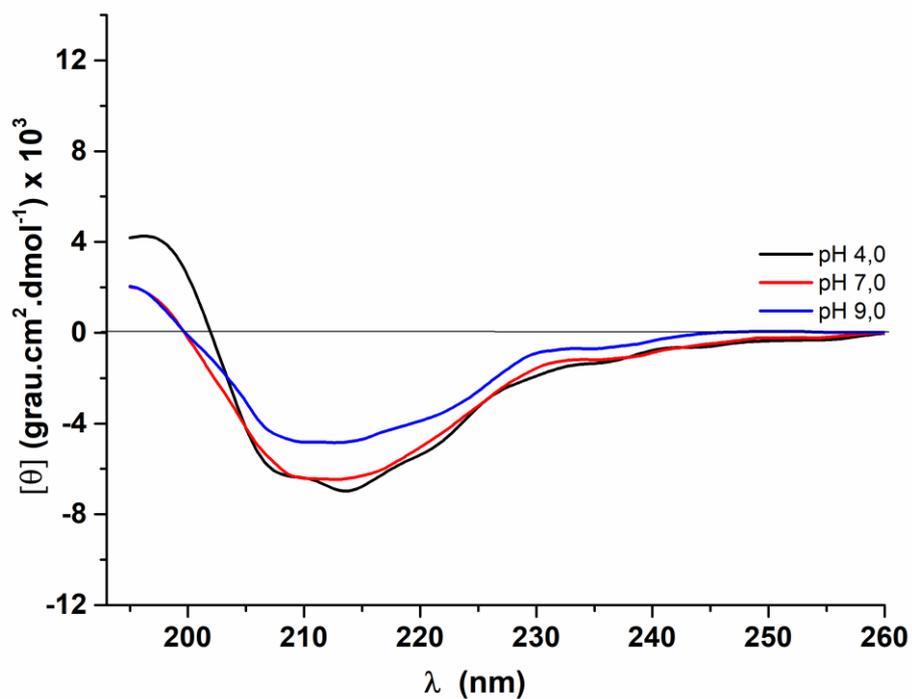


Figura 33 - Espectros dicróicos da proteína de 13 kDa na região distante da luz UV nos pHs 4,0; 7,0 e 9,0 obtidos a 25 °C. A concentração proteica foi 0,27 mg/mL em solução nos tampões acetato de sódio 5 mM para o pH 4,0 e Tris HCl 5 mM para os pHs 7,0 e 9,0.

Tabela 5 - Conteúdo de estrutura secundária da proteína de 13 kD presente na polpa de *E. dysenterica* obtido por deconvolução dos espectros de DC por meio do programa CDNN.

Estrutura Secundária (%)	pH 4,0	pH 7,0	pH 9,0
α-hélice	12,6	11,4	10,4
β-antiparalela	330,3	30,2	33,0
β-paralela	55,8	5,4	5,4
Volta-β	118,3	19,3	19,2
Desordenada	333,3	34,2	33,9

A termo estabilidade da proteína de 13 kDa foi analisada a partir das curvas de desnaturação com temperaturas variando entre 25 e 95 °C nos pHs 4,0, 7,0 e 9,0 (**Figura 34A**). As curvas mostram que nos três pHs estudados houve redução no sinal dicróico em 218 nm variando de -7.200 para -5.900 grau.cm²/dmol no pH 4,0; no pH 7,0 de -6.700 para -5.500 grau.cm²/dmol, já no pH 9,0 o sinal variou de -4.900 para -4.500 grau.cm²/dmol. Esse resultado caracteriza um processo de desnaturação proteica parcial. Os espectros dicróicos coletados concomitantemente em diferentes temperaturas, mostram que até 75 °C houve uma redução gradual do sinal dicróico em 208 nm, 218 nm e 222 nm para os três pHs analisados. No entanto, em 85 °C, a banda dicróica negativa deslocou-se para uma região próxima de 200 nm no pH 4,0. Já nos pHs 7,0 e 9,0 esse deslocamento só foi observado em 95 °C. Esse resultado caracteriza um aumento no teor de estrutura desordenada nas faixas de pH testadas (**Figura 34-D**).

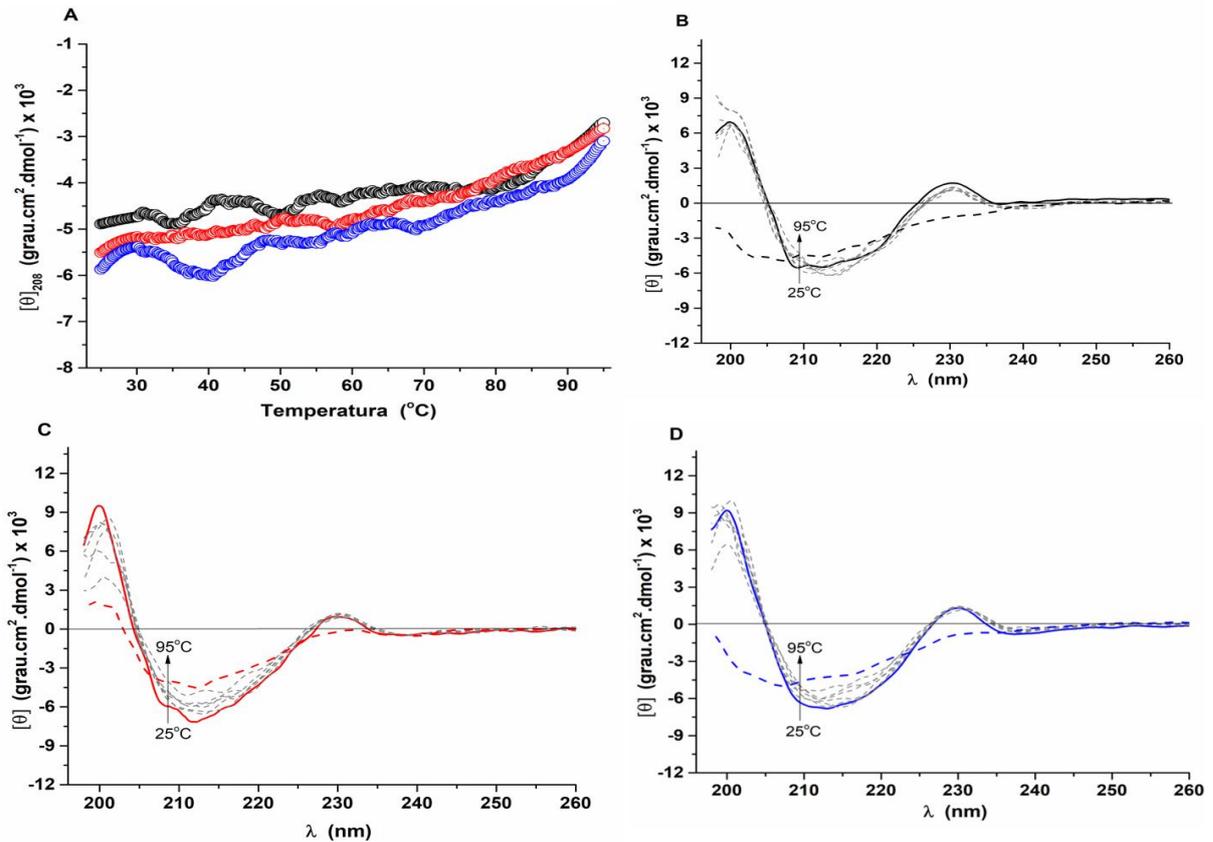


Figura 34 - Curvas de termoestabilidade e espectros dicróicos da proteína de 13 kDa em diferentes pH's. Curvas de termo estabilidade (A) e espectros dicróicos da proteína de 13 kDa presente na polpa de *E. dysenterica* nos pHs 4,0 (B), 7,0 (C) e 9,0 (D). As curvas de desnaturação térmica foram obtidas no comprimento de onda fixado em 218 nm, com variação da temperatura de 25 °C para 95 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul). Os espectros dicróicos da proteína foram obtidos na região distante da luz-UV de 25 °C a 95 °C, com intervalos de 10 °C (pH 4,0-linha preta, pH 7,0-linha vermelhas e pH 9,0-linha azul).

O perfil dos espectros dicróicos mostra que proteína de 13 kDa é mais termoestável em pH 9,0, pois houve uma redução menor do sinal a 218 nm variando somente em 400 grau.cm²/dmol. A proteína em estudo desenovelou-se parcialmente nos pHs testados, entretanto, não foi observado agregação. Portanto, um espectro dicróico coletado após ao ensaio de desnaturação ao retornar para 25 °C indicou que a proteína aparentemente se reenovelou nos três pHs testados, apresentando um espectro dicróico bem próximo ao espectro inicial obtido a 25 °C (**Figura 33**).

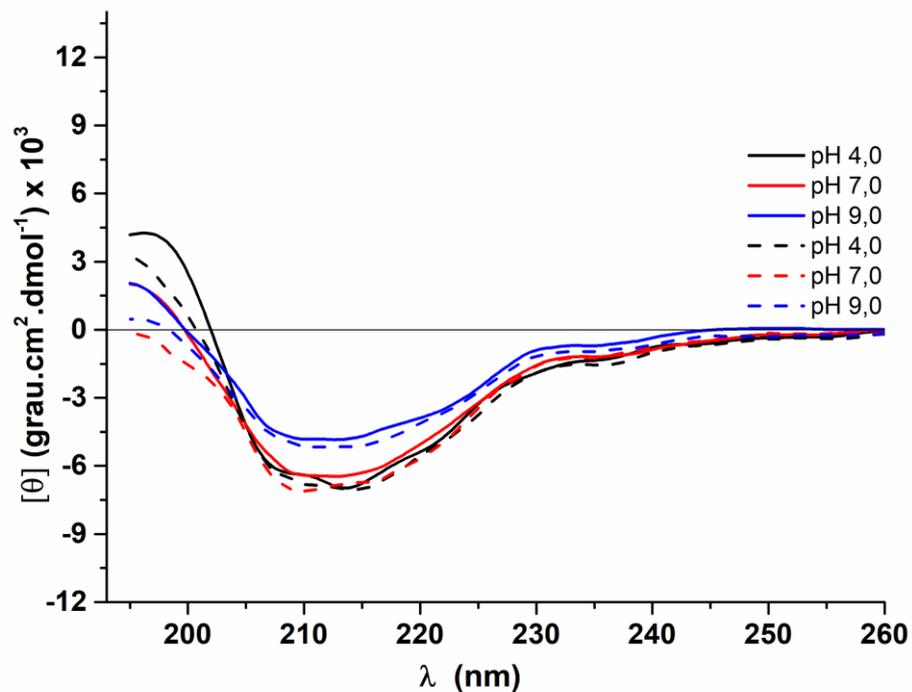


Figura 35 - Espectros dicróicos da proteína de 13 kDa durante e após ensaio de desnaturação. Espectros dicróicos da proteína de 13 kDa presente na polpa de *E. dysenterica* na região distante da luz UV nos pHs 4,0, pH 7,0 e pH 9,0 obtidos a 25 °C (linhas contínuas) e após ensaio de desnaturação térmica (25 a 95 °C) com retorno a 25 °C (linhas pontilhadas). Os espectros dicróicos da proteína de 9 kDa foram obtidos de 198-260 nm antes e após desnaturação térmica.

4 CONCLUSÃO

Eugenia dysenterica é uma espécie nativa do cerrado com elevado potencial econômico, que apresenta diversas aplicações (culinárias e medicinais) que são muito conhecidas pelos habitantes desta região. Mas, apesar dessa espécie ser bastante conhecida, poucos estudos em relação aos constituintes dessa planta, principalmente, em relação aos componentes químicos do fruto foram desenvolvidos.

Diante disso, a referida pesquisa teve como objetivo elucidar os constituintes da fração proteica, dos quais, três foram mais expressivos, com as seguintes massas: 7 kDa, 9 kDa e 13 kDa. O sequenciamento parcial de 38 resíduos de aminoácidos da proteína de 7 kDa, e o estudo de termo estabilidade e análise estrutural por meio da técnica de dicroísmo circular das três proteínas foram realizados e mostraram que estas apresentam boa resistência térmica o que é uma característica relevante para as futuras aplicações comerciais do fruto.

Os resultados desses estudos juntamente com a técnica de bioinformática de alinhamento de sequências, possibilitou a classificação da proteína de 7 kDa como uma nsLTP2 putativa. O que representa um avanço em relação aos estudos preliminares da fração proteica do fruto da cagaiteira.

No entanto, estudos adicionais são necessários para elucidação da fração proteica do fruto da cagaiteira e sobre suas funcionalidades por meio do sequenciamento total das proteínas, e estudos mais abrangentes que possibilitem a compreensão dos níveis de organização das proteínas presentes no fruto de *E. dysenterica*.

5 PESPERSCTIVAS

- Otimização do processo de extração.
- Sequenciamento total das proteínas presentes na polpa de *E. dysenterica*.
- Elucidação estrutural nos quatro níveis de organização biológicos.
- Testes de atividade biológicos.

6 REFERÊNCIAS

1. REIS, Amanda Figueiredo; SCHMIELE, Marcio. Características e potencialidades dos frutos do Cerrado na indústria de alimentos. *Brazilian Journal of Food Technology*, v. 22, 2019.
2. ALMEIDA, S. P.; SILVA, J. A.; RIBEIRO, J. F. Aproveitamento alimentar de espécies nativas dos cerrados: araticum, baru, cagaita e jatobá. Planaltina: EMBRAPA-CPAC, 1987. 83 p. EMBRAPA-CPAC. Documentos, 26.
3. CARVALHO, L.; et al. Study of volatile compounds in cagaita. 5th International Technical Symposium on Food Processing, Monitoring Technology in Bioprocesses and Food Quality Management. Potsdam: 1508-1512 p. 2009.
4. LIMA, T. B. et al. Identification of *E. dysenterica* laxative peptide: A novel strategy in the treatment of chronic constipation and irritable bowel syndrome. *Peptides*, New York, v. 31, p. 1426–1433, 2010.
5. SILVA, R.S.M.; CHAVES, L.J.; NAVES, R.V. Caracterização de frutos e árvores de cagaita (*Eugenia dysenterica* DC.) no sudeste do Estado de Goiás, Brasil. *Revista Brasileira de Fruticultura*, v.23, p.330-334, 2001.
6. SOUZA, E. R. B. et al. Fenologia de cagaiteira (*Eugenia dysenterica* DC.) no Estado de Goiás. *Revista Brasileira de Fruticultura, Jaboticabal*, v. 30, n. 4, p. 1009-1014, 2008.
7. DONADIO, L. C.; MÔRO, F. V.; SERVIDONE, A. A. Frutas brasileiras. Jaboticabal: Editora Novos Talentos, 2002. 288p.
8. LORENZI, H. Árvores brasileiras: Manual de identificação e cultivo de plantas arbóreas do Brasil. 4ª edição. Nova Odessa: Instituto Plantarum, 2002, v.1. p.368.
9. RIBEIRO, J. F.; SILVA, J. C. S.; BATMANIAN, G. J. Fitossociologia de tipos de cerrado em Planaltina - DF. *Revista Brasileira de Botânica*, São Paulo, v.8, p.131-142, 1985.
10. MARTINOTTO, C. Cultivo in vitro e aspectos morfofisiológicos de cagaiteira (*Eugenia dysenterica* DC.). 2004. 84 p. Dissertação (Mestrado em Agronomia/Fisiologia Vegetal) - Universidade Federal de Lavras, Lavras, 2004.
11. PRINCE, M.; WIMO, A.; GUERCHET, M.; ALI, G.-C.; WU, Y.-T.; PRINA, M. World alzheimer report 2015: The Global Impact of Dementia. London, p.87. 2015
12. CECÍLIO, A. B. et al. Screening of Brazilian medicinal plants for antiviral activity against rotavirus. *Journal of Ethnopharmacology [online]*. 2012, vol.141, n. 3. pp. 975-981.

13. DUARTE, A.; COSTA, A.; SANTOS, S.; FERRI, P.; PAULA, J.; NAVES, R. Changes in Volatile Constituents During Fruit Ripening of Wild *Eugenia dysenterica* DC. *Journal of Essential Oil Research*, v. 20, n. 1, p. 30-32, 2008
14. <http://carpoteca.blogspot.com/2013/12/cagaita-ou-cagaiteira.html> acesso em 02/01/2020
15. Arquivo pessoal
16. PRINCE, M.; WIMO, A.; GUERCHET, M.; ALI, G.-C.; WU, Y.-T.; PRINA, M. *World alzheimer report 2015: The Global Impact of Dementia*. London, p.87. 2015
17. CORRÊA, M.P. *Dicionário das plantas úteis do Brasil*. Rio de Janeiro: Imprensa Nacional, 1984. p.1926-1978.
18. SILVA, Cristian Aldemar Gasca. *Contribuição ao estudo químico e de atividade biológica de Eugenia dysenterica Mart. ex. DC. Berg (Myrtaceae)*. 2016
19. BRANDÃO, M. *Plantas medicamentosas do cerrado mineiro*. Informe Agropecuário, Belo Horizonte, v. 15, n. 168, p. 15-20, 1991.
20. SILVA, R. S. M. *Caracterização de subpopulações de cagaita (Eugenia dysenterica DC.) da região sudeste do estado de Goiás, Brasil*. 1999. 107 f. Dissertação (Mestrado) Universidade Federal de Goiás, Goiânia, 1999.
21. SILVA, M. M. M. *Estudo do desenvolvimento fisiológico da cagaita (Eugenia dysenterica)*. 2016. 119 f. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) - Universidade Federal de Goiás, Goiânia, 2016.
22. LIMA, T. B. et al. Identification of *E. dysenterica* laxative peptide: A novel strategy in the treatment of chronic constipation and irritable bowel syndrome. *Peptides*, New York, v. 31, p. 1426–1433, 2010.
23. NELSON, D. L.; COX, M. M. *Lehninger: Princípios de Bioquímica*. 6ª ed., Sarvier, 2014.
24. *Bioinformática da Biologia à flexibilidade molecular / organização de Hugo Verli*. - 1. ed. - São Paulo : SBBq, 2014. 282 p.
25. VENTURA M.M., XAVIER-FILHO J., A trypsin and chymotrypsin inhibitor from black-eyed pea (*Vigna sinensis*). I. Purification and partial characterization, *Ann. Acad. Bras. Cienc.* 38, 553–566, 1966

26. MENEGHETTI, Luisa. Proteassoma 20S de *Saccharomyces cerevisiae* em complexo com o inibidor de serinoproteases BTC1 de *Vigna unguiculata*. 2018. 98 f., il. Dissertação (Mestrado em Biologia Molecular)—Universidade de Brasília, Brasília, 2018
27. SOUZA, Amanda Araujo. Production and biochemistry characterization of the acid phosphatase *Trichoderma harzianum* (ALL42). 2011. 82 f. Dissertação (Mestrado em Ciências Biológicas) - Universidade Federal de Goiás, Goiânia, 2011.
28. FALICO, Diego Alejandro. Identificação, purificação e caracterização de peptídeos antimicrobianos presentes na secreção cutânea do anuro *Hypsiboas raniceps*. 2014. [13], 50 f., il. Dissertação (Mestrado em Biologia Animal)—Universidade de Brasília, Brasília, 2014
29. SOUZA, Amanda Araujo. Production and biochemistry characterization of the acid phosphatase *Trichoderma harzianum* (ALL42). 2011. 82 f. Dissertação (Mestrado em Ciências Biológicas) - Universidade Federal de Goiás, Goiânia, 2011.
30. FALICO, Diego Alejandro. Identificação, purificação e caracterização de peptídeos antimicrobianos presentes na secreção cutânea do anuro *Hypsiboas raniceps*. 2014. [13], 50 f., il. Dissertação (Mestrado em Biologia Animal)—Universidade de Brasília, Brasília, 2014
31. COSTA, Túlio OG et al. Isolation and amino acid sequencing by MALDI-TOF-MS/MS of a novel antimicrobial anionic peptide from the skin secretion of *Osteocephalus taurinus* (Anura, Hylidae). *Journal of the Brazilian Chemical Society*, v. 23, n. 12, p. 2133-2136, 2012
32. VIDAL, Julia Freitas Daltro. Caracterização biofísica de uma α 2-macroglobulina bacteriana e avaliação do seu potencial uso na identificação de proteases. 2018. 86 f., il. Dissertação (Mestrado em Biologia Molecular) — Universidade de Brasília, Brasília, 2018.
33. OLIVEIRA, Gideane Mendes de. Caracterização estrutural e enzimática da serino protease transmembrana II humana (TMPRSS2), um alvo potencial para o tratamento do câncer de próstata. 2018. 100 f. Dissertação (Mestrado) - Curso de Mestrado Me Biologia Molecular, Cel, Universidade de Brasilia, Brasilia - Df, 2018.
34. Niall HD (1973). «Automated Edman degradation: the protein sequenator». *Meth. Enzymol. Methods in Enzymology*. 27: 942–1010
35. Edman, P.; Högfeltdt, Erik; Sillén, Lars Gunnar; Kinell, Per-Olof (1950). «Method for determination of the amino acid sequence in peptides». *Acta Chem. Scand.* 4: 283–293.
36. Ashcroft, A. E. *An Introduction to Mass Spectrometry*. (2017)

37. MARTINS, Mariana Lima Boroni. Proteínas e peptídeos relacionados à defesa de plantas e análise proteômica de folhas de tomate inoculadas com *Xanthomonas Campestris* pv. vesicatoria. 2010. 94 f. Dissertação (Mestrado) - Curso de Pós-graduação em Bioquímica Agrícola, Universidade Federal de Viçosa, Minas Gerais, 2010.
38. van Loon, L. C. Induced resistance in plants and the role of pathogenesis-related proteins. *European Journal of Plant Pathology*, v.103, p.753-765. 1997.
39. Gozzo, F. Systemic acquired resistance in crop protection: From nature to a chemical approach. *Journal of Agricultural and Food Chemistry*, v.51, n.16, p.4487-4503. 2003.
40. Hancock, R. E. e Diamond, G. The role of cationic antimicrobial peptides in innate host defences. *Trends in microbiology*, v.8, n.9, p.402-410. 2000.
41. Yeaman, M. R. e Yount, N. Y. Mechanisms of antimicrobial peptide action and resistance. *Pharmacological reviews*, v.55, n.1, p.27-55. 2003
- Patterton, H.-G. Low energy peptide fragmentation pathways. *Adv. Biomol. Res. Clust.* 1–25 (2009).
42. Yeats, T. H. e Rose, J. K. C. The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein science : a publication of the Protein Society*, v.17, n.2, p.191-198. 2008.
- 28 Smith, P.K., et al. (1985). Measurement of protein using bicinchoninic acid. *Anal Biochem* 150:76-85
43. Bohlmann, H. e Apel, K. Isolation and characterization of cDNAs coding for leaf-specific thionins closely related to the endosperm-specific hordothionin of barley (*Hordeum vulgare* L.). *Molecular and General Genetics*, v.207, n.2-3, p.446-454. 1987.
- Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227: 680-685.
44. Balls, A. K. e Hale, W. S. A crystalline protein obtained from lipoprotein of wheat flour. *Cereal Chem*, v.19, p.279-281. 1942.
- MADDEN, Tom. The BLAST sequence analysis tool. In: McENTYRE, Jo; OSTELL, Jim (Org.). *The NCBI Handbook*. Bethesda: National Center for Biotechnology Information, 2002.
45. Bohlmann, H., et al. Leaf-specific thionins of barley a novel class of cell wall proteins toxic to plant-pathogenic fungi and possibly involved in the defence mechanism of plants. *The EMBO Journal*. 1998.
46. Thevissen, K., et al. Therapeutic potential of antifungal plant and insect defensins. *Drug discovery today*, v.12, n.21-22, p.966-971. 2007.
47. Wong, J. H., et al. A Review of Defensins of Diverse Origins. *Current Protein & Peptide Science: Bentham Science Publishers Ltd.* 8: 446-459 p. 2007.

48. Blein, J.-P., et al. From elicitors to lipid-transfer proteins: a new insight in cell signalling involved in plant defence mechanisms. *Trends in Plant Science*, v.7, n.7, p.293-296. 2002..
49. Yeats, T. H. e Rose, J. K. C. The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein science : a publication of the Protein Society*, v.17, n.2, p.191-198. 2008.
50. Kader, J. C. Proteins and the intracellular exchange of lipids. I. Stimulation of phospholipid exchange between mitochondria and microsomal fractions by proteins isolated from potato tuber. *Biochimica et biophysica acta*, v.380, n.1, p.31-44. 1975
51. Kader, J. C. Lipid-transfer proteins in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, v.47, n.1, p.627-654. 1996.
52. Boutrot, F., et al. Genome-wide analysis of the rice and Arabidopsis nonspecific lipid transfer protein (nsLtp) gene families and identification of wheat nsLtp genes by EST data mining. *BMC genomics*, v.9, p.86. 2008.
53. Finkina, E. I., et al. Purification and primary structure of novel lipid transfer proteins from germinated lentil (*lens culinaris*) seeds. *Biochemistry (Moscow)*, v.72, n.4, p.430-438. 2007.
54. Bakan, B. n. d., et al. Specific Adduction of Plant Lipid Transfer Protein by an Allene Oxide Generated by 9-Lipoxygenase and Allene Oxide Synthase. *Journal of Biological Chemistry*, v.281:, December 22, 2006, p.38981-38988. 2006.
55. Oshchepkova, Y., et al. Isolation of the lipid-transporting protein Ns-LTP1 from seeds of the garden fennel flower (*Nigella sativa*). *Russian Journal of Bioorganic Chemistry*, v.35, n.3, p.315-319. 2009
56. Pyee, J., et al. Identification of a Lipid Transfer Protein as the Major Protein in the Surface Wax of Broccoli (*Brassica oleracea*) Leaves. *Archives of Biochemistry and Biophysics*, v.311, n.2, p.460-468. 1994.
57. SCHÄGGER, Hermann; VON JAGOW, Gebhard. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical biochemistry*, v. 166, n. 2, p. 368-379, 1987Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227: 680-685.
58. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227: 680-685.
59. NESTERENKO, Michael V.; TILLEY, Michael; UPTON, Steve J. A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in

polyacrylamide gels. *Journal of biochemical and biophysical methods*, v. 28, n. 3, p. 239-242, 1994.

60. MADDEN, Tom. The BLAST sequence analysis tool. In: McENTYRE, Jo; OSTELL, Jim (Org.). *The NCBI Handbook*. Bethesda: National Center for Biotechnology Information, 2002.

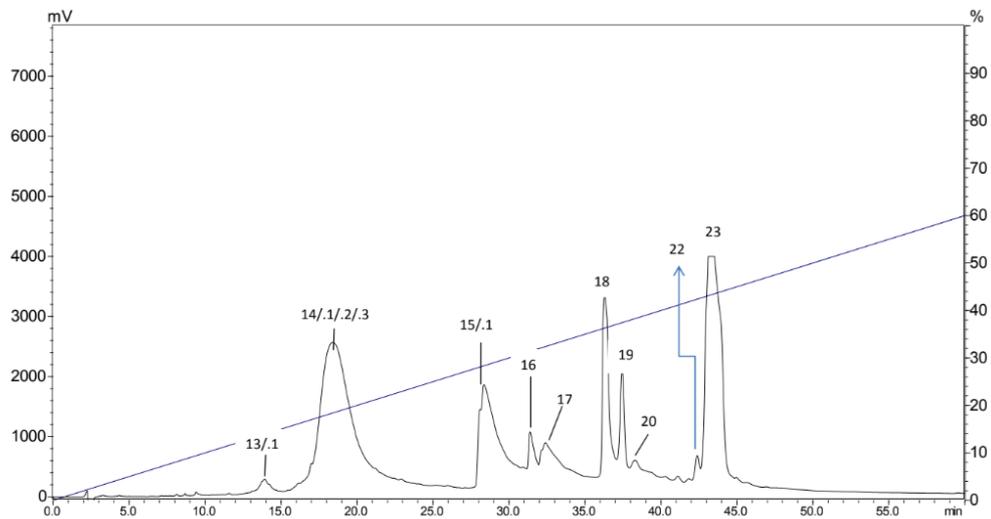
61. MARCHLER-BAUER, Aron et al. CDD: NCBI's conserved domain database. *Nucleic acids research*, v. 43, n. D1, p. D222-D226, 2014. Julie D. Thompson, Desmond G. Higgins, Toby J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Research*, Volume 22, Issue 22, 11 November 1994, Pages 4673–4680

62. WALLACE, B. A. Conformational changes by synchrotron radiation circular dichroism spectroscopy. *Nat. Struct. Biol.* 7, 708–709, 2000.

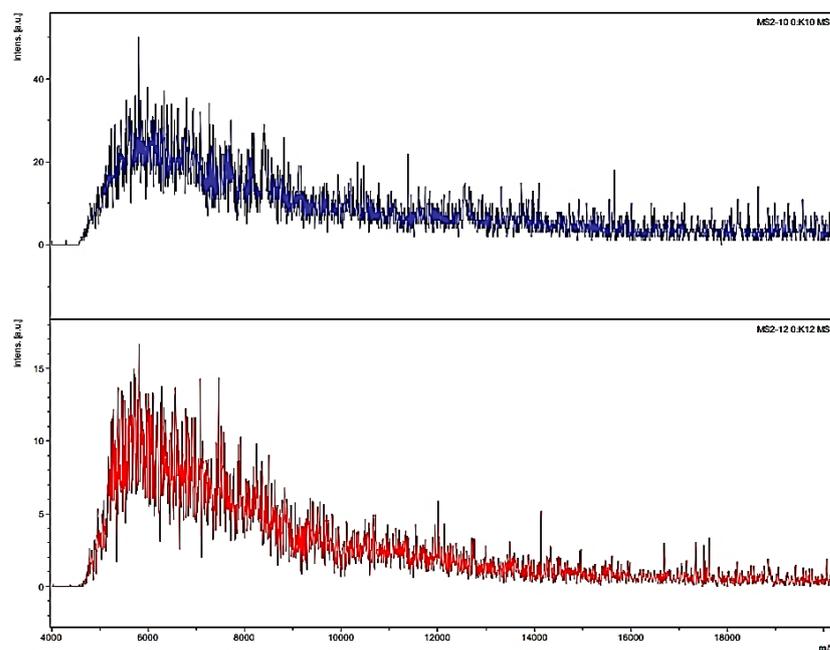
63. Smith, P.K., et al. (1985). Measurement of protein using bicinchoninic acid. *Anal Biochem* 150:76-85

7. ANEXOS

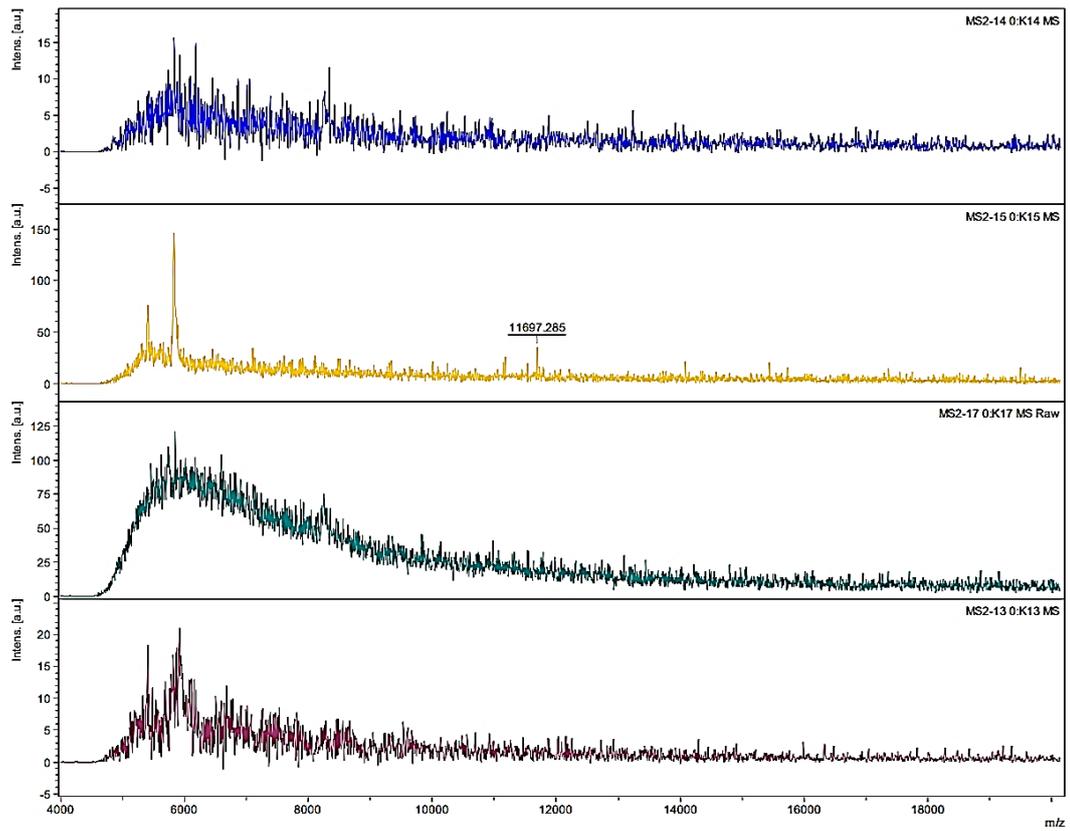
A seguir encontram - se o cromatograma da corrida de triagem para purificação da fração protéica do fruto da cagaiteira e os espectros de massas das frações coletadas.



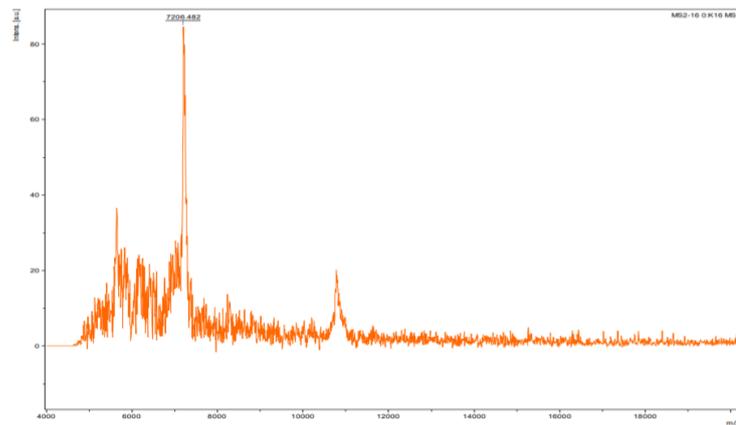
Cromatograma 1 - Corrida de triagem da fração proteica da cagaiteira foi usada uma coluna de fase reversa C18 (shimadzu VP ODS).



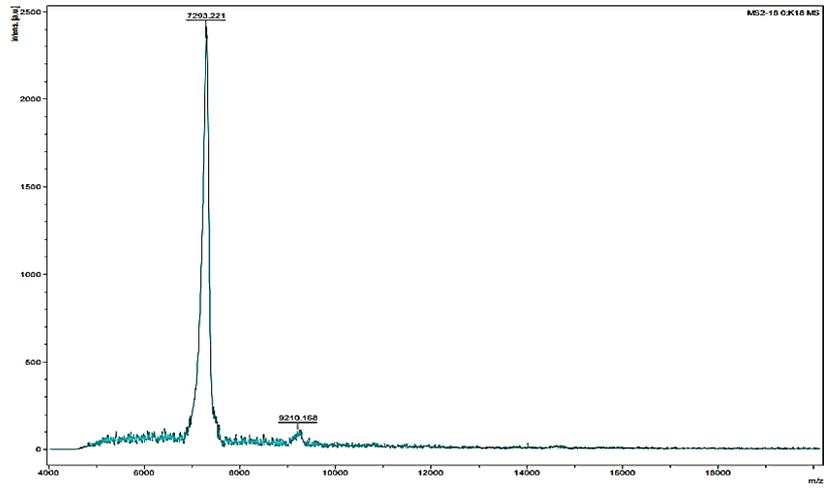
Espectros das frações 13 e 13.1 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaiteira



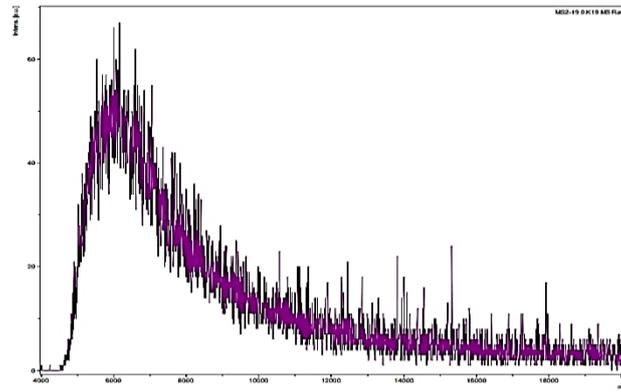
Espetros das frações 14.1/14.2/14.3/14 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaíta



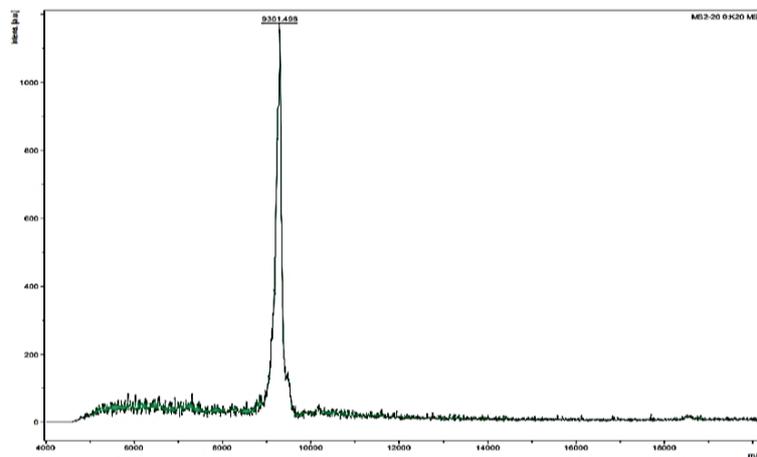
Espetro da fração 15 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaíta



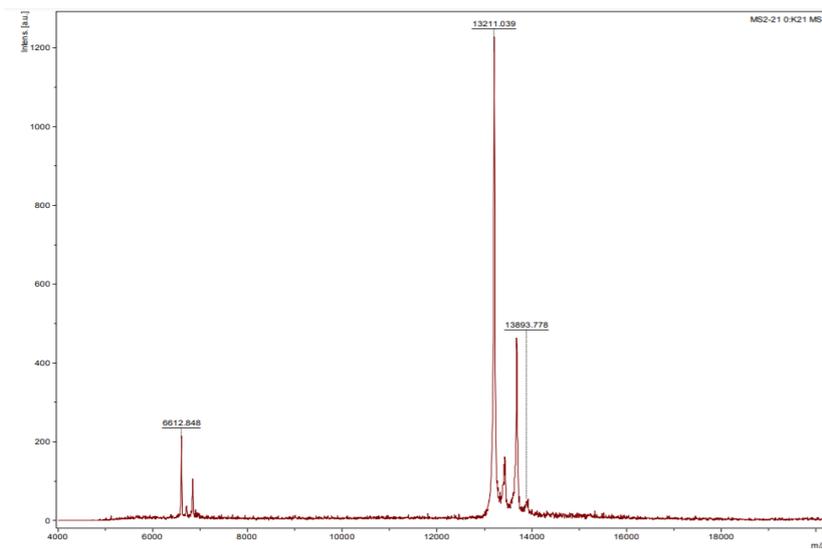
Espectro da fração 16 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaita



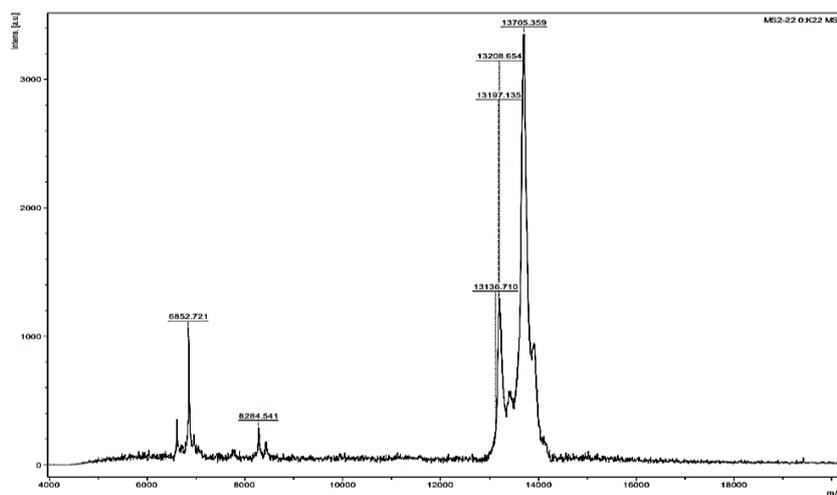
Espectro da fração 17 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaita



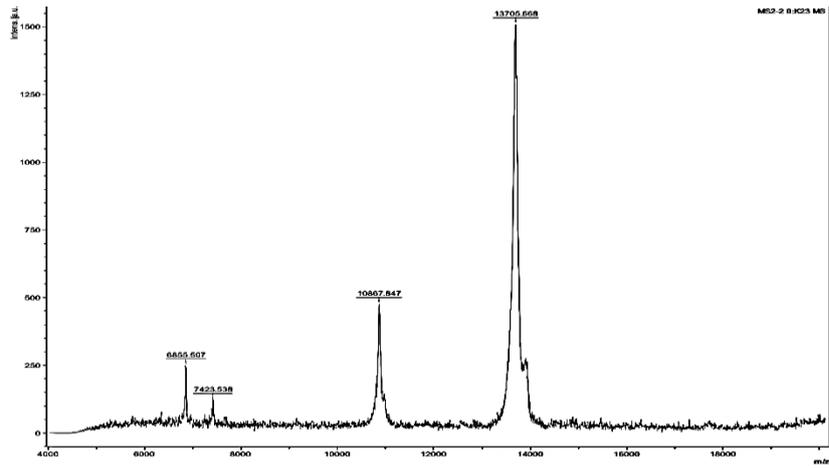
Espectro da fração 18 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaita



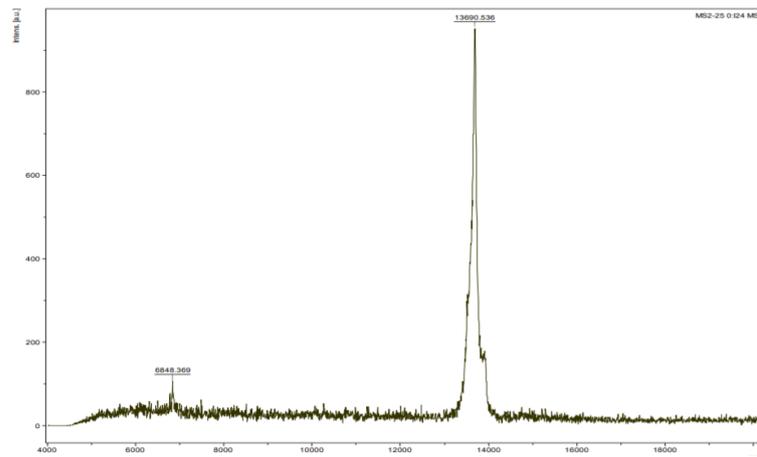
Espectro da fração 19 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaita



Espectro da fração 20 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaita



Espectro da fração 22 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaita



Espectro da fração 23 (4 a 20kDa) obtidos na corrida de triagem da fração proteica da cagaita