UNIVERSIDADE DE SÃO PAULO INSTITUTO DE FÍSICA DE SÃO CARLOS

JHON ANTONI VARGAS SANTILLAN

Caracterização estrutural e funcional de três enzimas com potencial uso biotecnológico da rota biossintética D-manose/L-galactose do ácido Lascórbico de Myrciaria dubia "camu-camu"

São Carlos

2023

## JHON ANTONI VARGAS SANTILLAN

Caracterização estrutural e funcional de três enzimas com potencial uso biotecnológico da rota biossintética D-manose/L-galactose do ácido Lascórbico de Myrciaria dubia "camu-camu"

> Dissertação apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

> Área de Concentração: Física Aplicada Opção: Física Biomolecular Orientador: Prof. Dr. Richard Charles Garratt

Versão Corrigida (Versão original disponível na Unidade que aloja o Programa)

> São Carlos 2023

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

> Vargas Santillan, Jhon Antoni Caracterização estrutural e funcional de três enzimas com potencial uso biotecnológico da rota biossintética Dmanose/L-galactose do ácido L-ascórbico de Myrciaria dubia (camu-camu) / Jhon Antoni Vargas Santillan; orientador Richard Charles Garratt - versão corrigida -- São Carlos, 2023. 122 p. Dissertação (Mestrado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2023. 1. Estrutura cristalográfica. 2. Biossíntese de vitamina C. 3. Cinética enzimática. 4. Espinafre. 5. Myrciaria dubia (camu-camu). I. Garratt, Richard Charles , orient. II. Título.

À minha mãe, Pilar Santillan, pelo amor infinito, compreensão e por sempre ter acreditado no meu potencial, que foi fundamental quando eu mais precisava. Amo você.

À minha família: irmãos, avós, tios e primos que sempre estiveram presentes durante minha formação.

À memória de meu pai Marlon Vargas, que com sua partida precoce deixou um grande vazio em nossos corações.

Á memória de minha tia Nonoy Santillan, sua partida deixa um grande vazio em nossos corações.

## AGRADECIMENTOS

A Deus, por me permitir continuar no caminho que sempre sonhei.

A minha família pelo apoio incondicional em toda minha formação profissional. Não há palavras que consigam expressar toda minha gratidão e amor por vocês. Em especial a minha mãe, Pilar Santillan, pelos conselhos e ânimos que apesar da distância sempre esteve perto.

Ao meu orientador, Prof. Dr. Richard Charles Garratt, por ter me aceitado como seu aluno, pela confiança em mim depositada, pelos conselhos e pelos gratos momentos compartilhados dentro e fora do laboratório. Muito obrigado Professor.

Ao Dr. Diego A. Leonardo Cabrejos pela paciência, por ter sido meu mentor, pelos conselhos e ensinamentos que me ajudaram a compreender um pouco sobre estruturas cristalográficas de proteínas ao longo de todos esses anos. Mas principalmente, por acreditar e incentivar o meu desenvolvimento intelectual, por ser um exemplo de pessoa, amigo, professor e pesquisador. Muito obrigado por tudo "Chitolin", sou eternamente grato.

À Dra. Susana pelos conselhos e dicas em muitos dos experimentos realizados neste trabalho. Além de se tornar minha amiga, sou eternamente grato Sú.

Ao Dr. Humberto Pereira, pelos conselhos e por estar sempre disposto a me ajudar nas análises enzimáticas e estruturais das proteínas envolvidas neste trabalho.

Aos técnicos do laboratório Biofísica e Biologia Estrutural que muito me ajudaram: Andressa, Bel e Rafael.

Aos meus amigos que conheci no laboratório, Leonardo (Leozinho), Tamiris, Gabriel, Higor, Italo, Danieli, Deborah, Adriano, Adriano Furtado (Jesús), Jakeline, Marco, Erick, Luis, Pamela e Stefany. Os dias foram melhores ao lado de vocês, sejam pelas horas de trabalho de laboratório, pelas sugestões, discussões, pelas piadas, os bares e os cafés.

Ao Dr. Juan Carlos Castro Gómez, pelas orientações iniciais e finais do projeto de pesquisa.

Ao meu amigo Eloy Condori, pela grata companhia nestes dois anos, pelo amparo e socorro em todos os momentos, o que fez com que nos tornássemos como irmãos.

A Grazieli pela compreensão e apoio em todo momento. Finalmente, a todos os novos amigos que conheci em São Carlos e meus amigos do Perú que sempre me demonstraram seu apoio.

Ao Instituto de Física de São Carlos (IFSC) pelas instalações e quadro técnico, que contribuíram para a minha formação científica.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 23038.011268/2020-05.

"Ciência e a vida cotidiana não podem e não devem ser separadas" Rosalind Franklin

"Si he logrado ver más lejos ha sido porque subido a hombros de gigantes" "Se consegui enxergar mais longe é porque procurei ver acima dos ombros dos gigantes." Isaac Newton

#### **RESUMO**

VARGAS, J. A. Caracterização estrutural e funcional de três enzimas com potencial uso biotecnológico da rota biossintética D-manose/L-galactose do ácido L-ascórbico de *Myrciaria dubia* "camu-camu". 2022. 122p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2022.

Myrciaria dubia (Kunth) McVaugh, conhecida como "camu-camu" ou "araçá d'agua", é uma fruta nativa da Amazônia, caracterizada por seu alto teor de L-Ácido Ascórbico (Vitamina C ou AsA). Devido a esta propriedade, as enzimas que participam da biossíntese de AsA poderiam ter potenciais aplicações biotecnológicos para a produção artificial de AsA. No entanto, a biologia estrutural das enzimas envolvidas tem sido bem pouco explorada. Aqui, descrevemos as propriedades biofísicas, funcionais e estruturais para três enzimas da via D-manose/Lgalactose (Smirnoff & Wheeler). As enzimas foram purificadas por cromatografia de afinidade e exclusão molecular. GDP-D-manose-3',5'-epimerase (MdGME), que catalisa uma dupla epímerização da GDP-manose para produzir GDP-L-galactose e GDP-L-gulose, provou ser um dímero em solução, suas estruturas cristalográficas foram resolvidas complexadas, sua forma apo (2,5 Å, ligada a NAD<sup>+</sup>) e holo (1.25 Å ligada a NAD<sup>+</sup>/substrato/produto). A L-galactose desidrogenase (MdGDH) e L-galactono-1,4-lactona desidrogenase (MdGalDH) provaram ser monoméricas em solução. MdGDH apresentou uma atividade catalítica com um Km de 0,128 mM e pH ótimo de 7. Demostramos que a inibição por AsA em MdGDH é devido a mudanças no pH e não necessariamente por uma inibição competitiva pelo sítio ativo como foi reportada para a enzima homóloga de espinafre (SoGDH). A estrutura cristalográfica de SoGDH foi resolvida a 1,4 e 1,75Å em suas formas apo e holo (ligado a NAD<sup>+</sup>), respectivamente. Finalmente a *Md*GalDH apresentou uma atividade catalítica com um K<sub>m</sub> de 0,044 mM e pH ótimo de 8 e apresentou uma inibição pelo próprio substrato. A estrutura cristalográfica foi resolvida a 2,1 Å. O presente trabalho contribui para uma compreensão mais amplo das relações estrutura-função das enzimas envolvidas na síntese da vitamina C.

Palavras-chave: Estrutura cristalográfica. Biossíntese de vitamina C. Cinética enzimática. Espinafre. *Myrciaria dubia* "camu-camu."

## ABSTRACT

VARGAS, J. A. Structural and functional characterization of three enzymes with potential biotechnological use from the D-mannose/L-galactose biosynthetic pathway of L-ascorbic acid from *Myrciaria dubia* "camu-camu". 2022. 122p. Dissertation (Master of Science) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2022

Myrciaria dubia (Kunth) McVaugh, known as "camu-camu" or "aracá d'agua", is a fruit native to the Amazon, characterized by its high content of L-Ascorbic Acid (Vitamin C or AsA). Due to this property, the enzymes that participate in the biosynthesis of AsA could have potential biotechnological applications for the artificial production of AsA. However, the structural biology of the enzymes involved has been poorly explored. Here, we describe the biophysical, functional, and structural properties for three enzymes of the D-mannose/L-galactose pathway (Smirnoff & Wheeler). The enzymes are purified by affinity chromatography and molecular GDP-D-mannose-3',5'-epimerase (MdGME), which catalyzes a double exclusion. epimerization of GDP-mannose to produce GDP-L-galactose and GDP-L-gulose, proved to be a dimer in solution, its crystallographic structure was solved, in both its apo (2.5 Å, bound to NAD<sup>+</sup>) and holo (1.25 Å bound to NAD<sup>+</sup>/substrate/product) form. L-galactose dehydrogenase (*Md*GDH) and L-galactono-1,4-lactone dehydrogenase (*Md*GalDH) proved to be monomeric in solution. *Md*GDH showed a catalytic activity with a K<sub>m</sub> of 0.128 mM and an optimal pH of 7. We show that inhibition by AsA in MdGDH is due to pH changes and is not necessarily due to a competitive inhibition at its active site as reported for the homologous enzyme of spinafre (SoGDH). The crystallographic structure of SoGDH was solved at 1.4 and 1.75Å in its apo and holo (NAD<sup>+</sup>-bound) forms, respectively. Finally, MdGalDH showed a catalytic activity with a K<sub>m</sub> of 0.044 mM and an optimal pH of 8 and was inhibited by its own substrate. The crystallographic structure was solved to 2.1 Å. The present work contributed to a broader understanding of the structure-function relationships of enzymes involved in the synthesis of vitamin C.

Keywords: Crystallographic structure. Biosynthesis of vitamin C. Enzymatic kinetics. Spinafre. *Myrciaria dubia* "camu-camu."

## LISTA DE FIGURAS

Figura 1 - Formas químicas do ácido ascórbico e análogos halogenados do AsA	24
Figura 2 - Vias metabólicas para biossíntese e reciclagem de vitamina C em <i>M. dubia</i>	28
Figura 3 - Fruto de Myrciaria dubia "camu-camu," fotoquímicos ativos <sup>84</sup> e seu uso na indústria	33
Figura 4 - Expressão e purificação de <i>Md</i> GME.	46
Figura 5 - Espectrometria de massa de <i>MdGME</i> .	46
Figura 6 - Cromatograma obtido através do SEC-MALS de <i>Md</i> GME	47
Figura 7 - Espectro de CD de <i>Md</i> GME	48
Figura 8 - Cristais da proteína de <i>Md</i> GME	49
Figura 9 - Topologia de <i>Md</i> GME	52
Figura 10 - Alinhamento de múltiplas sequências de quatro GMEs homólogos	53
<b>Figura 11 -</b> Esquema de ligação de NAD <sup>+</sup> de <i>Md</i> GME	54
<b>Figura 12 -</b> Diagrama estéreo de densidade de Fo-Fc não enviesada de GDP-α-D-manose (GD verde) e GDP-β-L-galactose (GDC, cor cinza).	D, cor 55
Figura 13 - Especificidade do substrato de <i>Md</i> GME.	56
Figura 14 - Sítio ativo de <i>Md</i> GME	57
Figura 15 - Mecanismo da GDP-manose 3',5'- epimerase.	58
Figura 16 - Estado fisiológico da enzima <i>Md</i> GME-NAD <sup>+</sup> .	60
Figura 17 - Os estados de abertura e fechamento de <i>Md</i> GME	61
Figura 18 - As interfaces da dimerização de <i>Md</i> GME ligado com o GDP-açúcar e NAD <sup>+</sup>	63
Figura 19 - Expressão e purificação de GDHs.	64
Figura 20 - Cromatograma de SEC-MALS de GDHs.	65
Figura 21 - Espectro de CD de GDHs.	67
Figura 22 - Propriedades cinéticas de camu-camu <i>Md</i> GDH e espinafre <i>So</i> GDH	68
Figura 23 - A inibição de <i>Md</i> GDH e <i>So</i> GDH por AsA em tampão Tris-HCl 100 mM (B e E) or HCl 300 mM (C e F) é mostrada	u Tris- 70
Figura 24 - Cristais das proteínas GDHs	73
Figura 25 - Topologia de <i>So</i> GDH	74
Figura 26 - Alinhamento de múltiplo das seguências entre GDHs e AKRs	76

Figura 27 - Rede de interações o	com o cofator NAD <sup>+</sup> da <i>So</i> GDH	77
Figura 28 - Mapa de <i>Polder</i> (a molécula NAD <sup>+</sup> da ca	a 2,5 σ) mostrando a qualidade do mapa de densidade eletrôn adeia A da estrutura <i>So</i> GDH-NAD <sup>+</sup>	ica da 77
Figura 29 - Interface de dimeriz redutase) comparada	zação em AKRs homodiméricos (representados por aflatoxina a com a estrutura de <i>So</i> GDH-NAD <sup>+</sup>	ıldeído 78
Figura 30 - O sítio ativo em SoC	GDH	79
Figura 31 - Comparação entre a SoGDH.	as interações de ligação de NADP <sup>+</sup> em AKRs e ligação de NA	D+ em 80
Figura 32 - Especificidade do su	ubstrato em AKRs e SoGDH	83
Figura 33 - O dendrograma m representantes de GD	ostra as relações topológicas entre alguns membros dos AKI PHs	Rs e 5 84
Figura 34 - Os estados de abertu	ara e fechamento do SoGDH	86
Figura 35 - Comparações da est	rutura GDH	88
Figura 36 - Mapa de densidade e SoGDH-NAD <sup>+</sup>	eletrônica (2Fo – Fc em 1 $\sigma$ ) de Loop- $\beta$ 3- $\alpha$ 1 e Loop- $\beta$ 4- $\alpha$ 2 em S	<i>o</i> GDH 88
Figura 37 - Expressão e purifica	ıção de <i>Md</i> GalDH	89
Figura 38 - Cromatograma de S	EC-MALS de <i>Md</i> GalDH	90
Figura 39 - Espectro de CD de l	MdGalDH	91
Figura 40 - Atividade de MdGal	IDH recombinante	92
Figura 41 - Cristais da proteína	do Kit PACT de <i>Md</i> GalDH	94
Figura 42 - Topologia de <i>Md</i> Ga	1DH	97
Figura 43 - Alinhamento de mu oxidase (ScAldO)	últiplas sequências de GalDH de outras espécies de plantas e	Alditol 100
Figura 44 - Sobreposição da enz	zima <i>Md</i> GalDH e <i>Bo</i> GalDH	101
Figura 45 - Cavidade de ligação	ao substrato xilitol de AldO e L-galL em <i>Md</i> GalDH.	104

## LISTA DE TABELAS

Tabela 1 - Propriedades gerais das proteínas estudadas	35
Tabela 2 - Coleta de dados de raios-X e estatísticas de refinamento de estrutura de MdGME	50
Tabela 3 - Resíduos envolvidos na dimerização de MdGME.	62
Tabela 4 - Parâmetros cinéticos para GDHs.	69
Tabela 5 - Medidas de pH no tampão de reação.	71
Tabela 6 - Coleta de dados de raios-X e estatísticas de refinamento de estrutura de GDHs.	72
Tabela 7 - Parâmetros cinéticos de estado estacionário para GalDH estudados até o momento.	93
Tabela 8 - Coleta de dados de raios-X e estatísticas de refinamento de estrutura de MdGalDH	95

## LISTA DE ABREVIATURAS E SIGLAS

Å	Ângstron
AF	AlphaFold2
AsA	Ácido Ascórbico
CD	Circular Dichroism - Dicroísmo circular
CDS	Coding DNA Sequence - Sequência de DNA codificante
HPLC	Cromatografia Líquida de Alta Eficiência
МАС	Immobilized Metal Affinity Chromatography - Cromatografia de
IMAC	afinidade por metal imobilizado
IPTG	Isopropil-β-D-tiogalactopiranosídeo
LB	Liquid Broth – Luria Bertani
MS	Espectrometria de Massa
NCBI	National Center for Biotechnology Information – Centro Nacional
	de Informação Biotecnologica
PDB	Protein Data Bank – Banco de Dados de Proteínas
SEC	Size Exclusion Chromatography - Cromatografia de Exclusão
	Molecular
SEC-MALS	Size Exclusion Chromatography - Multi Angle Light Scattering –
	Espalhamento de luz em múltiplos ângulos
SDS-PAGE	Dodecil-sulfato de Sódio de Poliacrilamida - Eletroforese em gel
	de poliacrilamida (PAGE) com Dodecil Sulfato de Sodio (SDS)
Tm	Temperatura de Melting

## SUMÁRIO

CAPÍ	TULO I	23
1	INTRODUÇÃO E OBJETIVOS	23
1.1	Aspectos gerais do Ácido L-Ascórbico (Ascorbato ou AsA)	23
1.2	Biossíntese do AsA	
1.3	Enzimas chaves da via D-manose/L-galactose	
1.3.1	GDP-manose-3',5'-epimerase (GME)	29
1.3.2	L-galactose desidrogenase (GDH)	30
1.3.3	L-galactono-1,4-lactona desidrogenase (GalDH)	
1.4	Vitamina C em Myrciaria dubia (Kunth) McVaugh (camu-camu)	31
1.5	Objetivos	34
1.5.1	Principal	34
1.5.2	Específicos	34
CAPÍ	TULO II	35
2	METODOLOGIA	35
2.1	Construção dos plasmídeos	35
2.2	Expressão e purificação das proteínas recombinantes	
2.3	Caracterização biofísica	
2.3.1	Determinação do estado oligomérico por cromatografia de exclusão	molecular
	acoplada ao espalhamento e luz a múltiplos ângulos (SEC-MALS)	
2.3.2	Análise do enovelamento e estabilidade térmica com espectroscopia de circular (CD)	dicroísmo 38
2.3.3	Análise de clivagem de <i>Md</i> GME	
2.3.4	Ensaios de Cristalização, coleta de dados e resolução das estruturas	
2.4	Caracterização bioquímica	41
2.4.1	Ensaios cinéticos GDHs	41
2.4.2	Ensaios cinéticos <i>Md</i> GalDH	
2.5	Análises de sequências homólogas	
2.6	Previsão da estrutura da proteína	44
CAPÍ	TULO III	45
3	RESULTADOS E DISCUSSÕES	45
3.1	Subcapítulo: GDP-mannose -3', 5'-epimerase (GME)	45
3.1.1	Purificação e determinação do estado oligomérico de <i>Md</i> GME	45

3.1.2	Enovelamento e estabilidade térmica de <i>Md</i> GME	47
3.1.3	Descrição geral da estrutura cristalográfica de <i>Md</i> GME	48
3.1.4	Estrutura de <i>Md</i> GME-NAD <sup>+</sup> -GDD/GDC	
3.1.5	Sítios de ligação do cofator NAD <sup>+</sup> e substrato GDP-D-manose	
3.1.6	Especificidade de ligação ao NAD <sup>+</sup>	54
3.1.7	Especificidade de ligação do GDP-açúcar (D-manose)	55
3.1.8	Mecanismo geral de GME	
3.1.9	Estrutura de <i>Md</i> GME-NAD <sup>+</sup>	
3.1.10	Estado aberto e fechado de <i>Md</i> GME	60
3.2	Subcapítulo: L-galactose desidrogenase (GDH)	63
3.2.1	O estado oligomérico da L-galactose desidrogenase	63
3.2.2	Enovelamento e estabilidade térmica de <i>Md</i> GDH e <i>So</i> GDH	66
3.2.3	Parâmetros cinéticos de GDHs	67
3.2.4	Descrição geral da estrutura de SoGDH	71
3.2.5	Ligação ao NAD <sup>+</sup>	75
3.2.6	Especificidade do substrato	
3.2.7	Estado aberto e fechado	
3.2.8	Comparação da estrutura GDH	
3.3	Subcapítulo: L-galactono-1,4-lactone desidrogenase (GalDH)	
3.3.1	Purificação e determinação do estado oligomérico de MdGalDH	
3.3.2	Enovelamento e estabilidade térmica de <i>Md</i> GalDH	
3.3.3	Parâmetros cinéticos de <i>Md</i> GalDH	91
3.3.4	Estrutura geral de <i>Md</i> GalDH	94
3.3.5	Sítio de ligação do cofator FAD	
3.3.6	Mecanismo geral de <i>Md</i> GalDH	
4	CONCLUSÃO	
	REFERÊNCIAS	

# **CAPÍTULO I**

## **1 INTRODUÇÃO E OBJETIVOS**

## 1.1 Aspectos gerais do Ácido L-Ascórbico (Ascorbato ou AsA)

Desde sua descoberta há 90 anos, o ácido L-ascórbico ou ascorbato (AsA, "*L-threo-hex-2enono-1,4-lactone*"), também chamado como vitamina C, <sup>1-2</sup> é um importante ácido orgânico que se caracteriza por atuar como agente redutor em mono e dioxigenases mantendo os metais em um estado reduzido. Além disso, participa dos fatores de liberação do hormônio do crescimento, corticotrofina e tirotropina, etc. <sup>2</sup> Como cofator atua nos processos de hidroxilação da biossíntese de carnitina e de resíduos dos aminoácidos de prolina e lisina. <sup>3</sup>

Estruturalmente o AsA apresenta um anel quase planar com cinco membros, com dois centros quirais que se resolvem nos quatro estereoisômeros, com uma massa molecular estimada de 176,12 g mol<sup>-1</sup>, densidade de 1,65 g cm<sup>-3</sup>, solúvel em água (> 30 g por 100 mL) e variavelmente solúvel na maioria dos álcoois e glicóis de cadeia curta (1-5 g por 100 ml).<sup>4</sup>

A propriedade química mais importante do ácido ascórbico é a sua oxidação reversível em ácido semideidroascórbico e em ácido desidroascórbico. Todas as ações conhecidas e postuladas até o momento são contabilizadas por uma única propriedade que é ser um doador de elétrons e, portanto, um agente redutor. <sup>4-5</sup> Sob condições fisiológicas >99% do AsA é encontrada na forma de ânion ascorbato, podendo doar sequencialmente dois elétrons da ligação dupla entre os carbonos dois e três. A perda do primeiro elétron (oxidação) produz o radical ascorbato (ácido semidehidroascórbico). Alguns radicais livres reativos produzidos por processos biológicos podem ser prejudiciais devido à sua natureza altamente reativa, porém estes podem ser reduzidos pelo ácido ascórbico, sendo este convertido (oxidado) no radical ascorbato. <sup>6</sup> Podendo, o radical ascorbato ser novamente AsA, alternativamente, ele também pode perder um segundo elétron (oxidado) para formar ácido desidroascórbico. <sup>4,6</sup> O ácido desidroascórbico é hidrolisado, com ruptura irreversível do anel para formar o ácido 2,3-dicetogulônico, cujos produtos incluem oxalato, treonato e possivelmente xilose, ácido xilônico e ácido linxônico. Análogos halogenados da vitamina C podem ser oxidados em ácido bromo desidroascórbico e ácido iodo desidroascórbico (**Figura 1**). <sup>4</sup>



Figura 1- Formas químicas do ácido ascórbico e análogos halogenados do AsA.

Fonte: Adaptada de PADAYATTY et al.<sup>4</sup>

O AsA pode atuar como um sequestrador de radicais livres em reações envolvendo espécies reativas de exigênio (ROS), eliminando facilmente espécies reativas de oxigênio e nitrogênio (superóxido, hidroperoxila, peroxinitrito e nitróxido), além de estar associado com a proteção de lipídios, ácidos nucleicos e proteínas contra oxidantes. <sup>7-8</sup>

Por outro lado, o metabolismo do AsA é muito complexo. O modo de ingresso às células, que é dependente do Na<sup>+</sup>, e seu acúmulo celular é devido aos sistemas celulares de redução de dehidroascorbato (DHA). O sistema enzimático acessório necessário para reduzir o potencial de acumulação transitória do radical ascorbato, fez com que os organismos evoluíssem para manter o AsA em sua forma reduzida. Nas plantas, sabe-se que a desidroascorbato redutase (DHAR) catalisa a regeneração do ácido ascórbico a partir do seu estado oxidado e serve como um importante regulador da reciclagem do AsA. Em células animais, DHAR também é reconhecido por sua atividade intrínseca de tioltransferases (glutarredoxinas) e proteína dissulfeto isomerase. <sup>5</sup>

Como efeito, o AsA é um dos nutrientes que possui maior importância nos sistemas essenciais aos seres vivos e graças a sua capacidade antioxidante no homem, faz com que seu mecanismo de ação seja de importância global nos diversos processos fisiológicos, imunológicos e biológicos. O resultado é proporcionar a melhoria em algumas doenças, como por exemplo seu efeito na geração de respostas imunológicas, a epigenética, a diabetes, o escorbuto e recentemente seu efeito positivo na doença do Alzheimer.<sup>9</sup>

Além disso, o AsA também apresenta um importante papel na agricultura, bioindústria e nutrição, sendo sintetizada como suplementos vitamínicos, preparações farmacêuticas, cosméticos (estimulante para a produção de colágeno) e atualmente exploradas no processamento de diversos alimentos. <sup>10–13</sup> Atualmente, a maior parte do AsA fabricada comercialmente é sintetizada através do processo químico de fermentação de Reichstein & Grussner. <sup>14</sup> A fabricação do AsA baseia-se na conversão D-glicose a AsA mediante uma série de etapas químicas e uma fermentação bacteriana para a conversão de D-sorbitol em L-sorbose. <sup>10</sup> No geral, o método apresenta a vantagem de ter uma aplicabilidade de desenvolvimento, mas com sérias limitações enquanto a sua baixa eficiência, alta demanda energética e sobretudo pelos danos ambientais que causa com o uso dos diversos insumos químicos (acetona, ácido sulfúrico e hidróxido de sódio) que fazem parte do processo de síntese. <sup>10</sup> Esses e outros fatores econômicos geraram um interesse substancial na exploração de biotransformações microbianas na fabricação de AsA. <sup>10</sup>

Os avanços na bioquímica, tecnologia do DNA recombinante e a revolução genômica acrescentaram métodos alternativos com o objetivo de minimizar as limitações do método de Reichstein. Processos de fermentação foram melhorados com a adição exógena de cofatores ou nutrientes, <sup>15-16</sup> de reconstrução dos microrganismos auxiliares, <sup>16-17</sup> de otimização da via metabólica dos microrganismos de produção, <sup>18–20</sup> de desenhos de novos sistemas de cocultura de linhagens auxiliares, <sup>21-22</sup> mas nenhum conseguiu resolver ou suplantar as limitações do método de Reichtein & Grussner.

Dada a importância e aplicabilidade do AsA nos diferentes processos agrícolas, industriais e farmacêuticos, esforços recentes foram desenvolvidos com ajuda da engenharia de proteínas, engenharia metabólica e das estratégias de biologia sintética. <sup>16,23</sup> Trabalhos como a construção de linhagens capazes de sintetizar AsA foi reconstruída em *Sacharomyces cerevisiae*. Cinco genes exógenos de *Arabidopsis* foram introduzidos nas células do chassi (máquinas mínimas auto-replicantes) para obter a cepa YLAA, encarregada de abrigar a

biossíntese de AsA, <sup>16,24</sup> mas a produção continua sendo baixo. De forma semelhante, em *Kluyveromyces lactis*, realizou-se a construção de plasmídeos modificados mediante a integração dos genes da via da L-galactose de *A. thaliana*, para avaliar um processo de fermentação sem usar precursores de isómeros "L" no meio de cultura, resultando na produção de AsA na ausência de quaisquer precursores de AsA como intermediários de L-galactose, L-galactono-1,4-lactona ou L-gulono-1,4-lactona no meio de crescimento. No entanto, a produção com os precursores naturais (exemplo: L-galactose) foi 3 vezes maior, o que resulta ainda em um elevado custo de produção. <sup>25</sup>

Por outro lado, a síntese em grande escala do AsA por microalgas ainda apresenta um elevado custo, devido ao cultivo que é usado na produção, sendo superior ao cultivo dos microrganismos. <sup>26</sup> Portanto, há necessidade de desenvolver um processo de biossíntese eficaz de AsA, acessível e amigável com o meio ambiente. Avanços recentes sobre a biossíntese em plantas poderiam ser usados para direcionar os esforços na compreensão dos mecanismos subjacentes a este processo complexo. Portanto, estudos sob condições que estimulam a produção e acúmulo de AsA são necessários para fornecer uma visão biotecnológica da produção do AsA a nível industrial. <sup>23</sup>

## 1.2 Biossíntese do AsA

A maioria dos animais (anfíbios, répteis, aves e mamíferos) são capazes de sintetizar o AsA a partir da glicose no rim ou fígado. No entanto, em muitos invertebrados, peixes teleósteos, algumas espécies de mamíferos (morcegos frugívoros) e na maioria dos primatas, incluindo o homem, perderam a capacidade de sintetizar o AsA, por possuírem um gene não funcional. <sup>4,14,27-28</sup> Esta deficiência foi localizada pelo acúmulo de várias mutações no gene que codifica, L-gulono-1,4-lactona oxidase (GuLO, EC. 1.1.3.8), a última enzima que catalisa a etapa terminal na síntese do AsA. Em animais capazes de sintetizar o AsA, a biossíntese ocorre apenas em alguns tipos celulares, chegando a ser muitas vezes insuficiente. Portanto, os alimentos vegetais são a principal fonte deste micronutriente essencial para os organismos vivos, especialmente para os humanos. <sup>1,4-5,7</sup> Na maioria dos mamíferos, quando o AsA é consumido na dieta, o íleo e o jejuno são os principais locais de absorção. As necessidades são atendidas com uma ingestão de 20 a 80 mg por 1.000 kcal, mas grande parte da dose absorvida (>50%) não é metabolizada. <sup>5</sup>

Atualmente, nas plantas, existem múltiplas rotas biossínteticas para a produção do AsA.<sup>29-32</sup> embora ainda não seja claro quantas destas são relevantes fisiologicamente. Até agora a via D-manose/L-galactose é a mais estudada, <sup>33</sup> conhecida também como a via de Smirnoff-Wheeler (SW), e a mais respaldada por estudos bioquímicos e genéticos  $^{27-28,34}$  (Figura 2). A D-manose-1-fosfato é convertida em AsA mediante seis enzimas, a GDP-D-manose pirofosforilase (GMP, E.C. 2.7.7.13) que converte a D-manose-1fosfato em GDP-D-manose, que subsequentemente é epimerizada pela GDP-α-D-manose-1',5'-epimerase (GME, em GDP-L-galactose. GDP-α-D-manose e GDP-L-galactose E.C.5.1.3.18) são adicionalmente usadas na síntese de polissacarídeos da parede celular e glicosilação de proteínas, interconectando assim a via SW com a biossíntese da parede celular.<sup>27</sup> O próximo passo é catalisado pela GDP-L-galactose fosforilase (GGP, E.C. 2.7.7.69), produzindo Lgalactose-1-fosfato. Esta etapa tem sido postulada como ponto de controle da via. <sup>27,34-35</sup> Em seguida, L-galactose 1-fosfato fosfatase (GPP, E.C. 3.1.3.25) e outras fosfatases não identificadas, 27,36-37 transformam L-galactose 1-fosfato em L-galactose. A L-galactose (Lgal) é então oxidada a L-galactono-1,4-lactona (L-galL) pela L-galactose desidrogenase (GDH, E.C.1.1.1.3.16). Acredita-se que todas essas etapas ocorram no citosol, enquanto a Lgalactono -1,4-lactona atravessa a membrana externa da mitocôndria para ser oxidada pela Lgalactono 1.4-lactona desidrogenase (GalDH, E.C. 1.3.2.3), componente integral do complexo 1 de transporte de elétrons mitocondrial na membrana interna das plantas.<sup>38</sup>



Figura 2 - Vias metabólicas para biossíntese e reciclagem de vitamina C em *M. dubia*. A via Biossintética de D-manose/L-galactose (SW) está indicada em azul: As etapas catalisadas por GDP-manose-3',5'-epimerase (GME), L-galactose desidrogenase (GDH), L-galactona-1,4-lactona desidrogenase (GalDH) está indicada em vermelho. mIM= membrana interna da mitocôndria, mIMS= Espaço intermembrana da mitocôndria, mOM= membrana externa da mitocôndria. Os nomes das demais vias estão indicados em cores diferentes.

Fonte: Elaborada pelo autor.

#### 1.3 Enzimas chaves da via D-manose/L-galactose

Até hoje os esforços foram aplicados na compreensão da biossíntese da vitamina C através das enzimas que participam na via de SW, os quais foram o foco de estudo nos últimos 20 anos. No momento existem no mínimo três postulados como enzimas chaves na regulação da

via SW, GME, GDH e GalDH. Os papéis de GME, GDH e GalDH em sua maioria foram relatados a partir de *Arabidopsis thaliana*, *Oriza sativa* e *Actinidia deliciosa (kiwi)*, comumente utilizadas como organismos modelos. Estas espécies de plantas são diferentes enquanto a suas características bioquímicas, fisiológicas e genéticas, mais diferente ainda na produção de AsA.

### 1.3.1 GDP-manose-3',5'-epimerase (GME)

A primeira atividade reportada para GME foi há 55 anos nos organismos de *Helix pomatia* e *Chorella pyrenoidosa*, <sup>39–41</sup> mas somente no início do século 21 que estudos de caracterização bioquímica para plantas, foram relatadas para a planta modelo de *Arabidopsis thaliana* (*At*). <sup>42-43</sup> A identificação de dois produtos da epimerização da GDP-D-manosse catalisados pelo GME levou a propor uma via alternativa de biossínteses de AsA em plantas <sup>33,44</sup>. No entanto, GDP-D-manose além de ser o substrato do segundo passo da biossíntese de AsA, é também um precursor conhecido da síntese de D-manose, L-fucose e L-galactose e, portanto, também da pectina ramnogalacturonan-II (RG-II) e das hemiceluloses. <sup>45</sup>

A interconexão entre a biossínteses de AsA e do polissacarídeo da parede celular, faz com que o GME seja um regulador chave destes mecanismos. <sup>46–49</sup> No entanto, as evidências experimentais em apoio a essas hipóteses são limitadas. Atualmente, estudos de expressão em GME mostraram relação significativa com o aumento do conteúdo de AsA nas plantas, <sup>50–53</sup> sugerindo uma estratégia de regulação do conteúdo de AsA para aliviar o estresse abiótico. <sup>47</sup> GME é a proteína mais conservada da biossíntese de AsA, <sup>44,54</sup> e as análises evolutivos revelaram que o gene de GME pode possuir mais de um transcrito em algumas plantas. <sup>46,48</sup>

GME de *Arabidopsis thaliana* foi reportada como a primeira estrutura cristalográfica da via de SW faz 17 anos. <sup>55</sup> Após resolvida a estrutura, estudos de caracterização bioquímica e fisiológicas foram realizados em outras espécies de plantas, como, *Oryza Sativa* L,<sup>56</sup> *Solanum lycopersicum*, <sup>49,57</sup> e variantes de kiwi (*Actinidia rufa* e *Actinidia deliciosa*). <sup>58</sup> Estudos recentes detectaram a produção de GDP-Altrose pelo GME, este novo nucleotídeo-açúcar seria o terceiro produto de GME, enzima extremamente especial e interessante, pelo fato dela produzir três produtos em uma mesma reação. Atualmente as epímerases em geral são de suma importância para a síntese de açucares raros. Estudos feitos em *Methylacidiphilum fumariolicum* cepa solV(organismo termofílico) a caracterizaram como um biocatalizador industrial para a produção de L-açúcares e derivados. <sup>59-60</sup> Neste sentido, informações

estruturais de outras GMEs precisam ser desenvolvidas para um melhor entendimento e aplicação industrial desta enzima, faz quase duas décadas desde que a primeira e última estrutura foi resolvida. <sup>41,55</sup>

#### 1.3.2 L-galactose desidrogenase (GDH)

A GDH catalisa a oxidação da L-galactose (L-gal) na posição C1 para produzir Lgalactono-1,4-lactona (L-galL) na penúltima etapa da síntese de AsA pela via SW. Esta enzima é reconhecida como uma das mais importantes na regulação da biossíntese de AsA. <sup>33,61-62</sup> Estudos bioquímicos foram realizados a partir de folhas ou expressão heteróloga em diferentes espécies de plantas, como L-galactose desidrogenase de *Pisum sativum (Ps*GDH), *Arabidopsis thaliana (At*GDH), *Spinacia oleracea (So*GDH) e *Actinidia deliciosa (Ad*GDH) <sup>61-63</sup>. Com base nestes estudos, demonstraram que GDH possui uma alta afinidade e seletividade para L-gal com um K<sub>m</sub> variando de 0,08-0,43 mM e com preferência por NAD<sup>+</sup> no lugar de NADP<sup>+</sup> como seu cofator. <sup>61-63</sup> Além disso, o AsA demonstrou-se inativar lenta e irreversivelmente a enzima GDH, atuando como um inibidor de feedback através da inibição competitiva com o substrato. Isso sugere que a GDH pode desempenhar um papel significativo na regulação da biossíntese de AsA em plantas. <sup>62-63</sup> Os estudos estruturais desta enzima foram limitados, porém, recentemente os investigadores Momma & Fujimoto disponibilizaram duas estruturas no PDB (7EZL e 7EZI) de GDH de *Oriza sativa (Os*GDH), <sup>64</sup> mas as características estruturais de *Os*GDH ainda não foram descritas na literatura.

### 1.3.3 L-galactono-1,4-lactona desidrogenase (GalDH)

A etapa final da biossíntese de AsA exige a oxidação da L-galL a AsA, em uma reação que é catalisada pela L-galactono-1,4-lactone desidrogenase (GalDH). Essa reação não é acoplada a nenhum par de coenzimas, de modo que media a oxidação de dois elétrons de L-galL em AsA com a redução concomitante de Citocromo C (Cc). <sup>65</sup> GalDH é uma flavoenzima de membrana integral da membrana mitocondrial interna onde transporta elétrons para a cadeia de transporte de elétrons via Cc. <sup>66–68</sup> GalDH e outras aldonolactonas oxidorreductase pertencem à família de flavoproteínas vanilil-álcool oxidase (VAO). <sup>69</sup> Os membros desta família compartilham uma topologia de enovelamento de dois domínios compreendendo um domínio de ligação FAD N-terminal conservado e um domínio Cap C-terminal que determina a especificidade do substrato. O sítio ativo está localizado na interface dos domínios. <sup>70-71</sup> A maioria das aldonolactonas oxidorreductases são oxidases produtoras de peróxido de

hidrogênio contendo FAD ligado covalentemente, enquanto GalDH reage mal com oxigênio molecular e contém FAD ligado de forma não covalente.<sup>72</sup> Além de produzir L-ascorbato, GalDH é importante para o funcionamento correto das mitocôndrias vegetais, <sup>73</sup> e necessária para o montagem do complexo respiratório I das plantas.<sup>38,65,67</sup> A GalDH foi estudada em várias espécies de plantas como por exemplo: Solanum lycopersicum (SlGalDH), 73 Arabidopsis thaliana (AtGalDH), <sup>69,72,74-75</sup> Triticum aestivum (TaGalDH), <sup>68,76</sup> Ipomoea batatas L., cv. Kintoki (IbGalDH), 77 Brassica oleracea, var. botrytis (BoGalDH), 78 Nicotiana sylvestris (NiGalDH),<sup>67</sup> e Capsicum annuum L (CaGalDH).<sup>25</sup> O produto primário de tradução de GalDH tem uma massa molecular de aproximadamente 68 kDa (possui um sinal peptídeo), mas é processado em uma proteína madura de 56 a 58 kDa.<sup>65</sup> GalDH mostrou uma alta afinidade pelo seu substrato natural L-galactono-1,4- lactona (L-galL), e baixa afinidade pelo substrato análogo L-gulono-1,4- lactona (L-gulL). Recentemente a única estrutura de GalDH foi resolvida como parte do complexo respiratório I (complexo I) de Brassica oleracea, <sup>38</sup> mas com uma baixa resolução (3.8 Å). Embora nos últimos 10 anos os estudos realizados em GalDH a partir de várias fontes de plantas tenham sido muito relevantes, informações como as diferenças na ligação do cofator, a especificidade pelo substrato e os resíduos importantes do sítio ativo ainda não foram explorados. Além disso, ainda não há na literatura uma descrição detalhada da estrutura 3D de GalDH de qualquer espécie.

#### **1.4** Vitamina C em Myrciaria dubia (Kunth) McVaugh (camu-camu)

*Myrciaria dubia* (Kunth) McVaugh, também conhecida como "camu-camu" ou "araçá d' agua", é um típico arbusto frutífero nativo encontrado nas florestas tropicais da Amazônia. <sup>79</sup> Em seu ambiente nativo, a árvore pode atingir de 4 a 8 m de altura, podendo permanecer até seis meses submersa na água, e cresce naturalmente na área de várzea. Além da região amazônica, a espécie é encontrada nos estados do Pará e Rondônia (Brasil). <sup>80-81</sup> Os frutos são globulares, medem de 1,0 a 3,2 cm de diâmetro, pesa de 2 a 20 g, possuem de uma a quatro sementes por fruto e 50 a 55% de polpa branca com casca fina e brilhante, que varia do rosa ao vermelho escuro ou mesmo roxo escuro quando é totalmente maduro. <sup>80,82-83</sup> Populações naturais desta espécie crescem em áreas densas expostas a inundações de córregos e nas margens de rios, lagos ou pântanos da região amazônica (Perú, Brasil, Colômbia, Equador e Venezuela), <sup>80,82</sup> mas a maior concentração das populações naturais e fonte de variabilidade genética, é encontrada na região de Loreto da Amazônia peruana. <sup>79</sup> Várias partes da planta são usadas na medicina tradicional no tratamento de doenças agudas e crônicas,<sup>84</sup> por

exemplo: asma, arterosclerose, catarata, depressão, gripe, gengivite, glaucoma, hepatite, infertilidade, enxaqueca, osteoporose, mal de Parkinson, anti-inflamatório, diabetes e malária. <sup>79,85–87</sup> Alguns de esses atributos benéficos são bem apoiados por estudos *in vitro e in vivo*, bem como ensaios clínicos em humanos. <sup>84</sup>

Geneticamente, a planta é caracterizada por um genoma diplóide com diversidade genética moderada, apresenta uma notável capacidade de biossintetizar e acumular quantidades importantes de uma variedade de fitoquímicas promotores de saúde, especialmente antioxidantes, o que tornou os frutos de camu-camu de interesse para as indústrias farmacêuticas, cosmética e alimentícia. <sup>84,88</sup> Sua importância econômica mais sobressaliente está no alto teor de vitamina C nas frutas, o qual varia entre tecidos e estados de maduração do fruto, com uma concentração maior na casca entre 0,96-3,1 por g/100 g de polpa. <sup>79,82-83,89</sup> Em comparação, as espécies de plantas consideradas referências de teor de vitamina C, como laranja ou kiwi, apresentam concentrações de apenas 0,05 g e 0,085 g/100 g de polpa respectivamente, <sup>90</sup> é por isso que camu-camu é considerada uma das mais ricas fontes de AsA de todas as espécies vegetais. <sup>91</sup> Além do AsA e antocianinas, pelo menos 30 flavonóis glicosídeos (polifenóis em torno de 739 mg.100<sup>-1</sup>g de peso corporal) parecem contribuir para a capacidade antioxidante global da polpa do fruto de camu-camu. <sup>80,92</sup>

Suas propriedades e usos o tornaram uma espécie de planta muita valiosa para a economia. A fruta é consumida naturalmente ou processado (polpa, sucos, geleias, etc.) e precisamente a polpa congelada é exportada para o Japão, Ásia, Europa e Estados Unidos, sendo utilizado como principal produto de valor agregado para as diferentes indústrias já mencionadas acima <sup>79-80</sup> (**Figura 3**). Na indústria alimentícia é comumente usado para bebidas mistas para aumentar o teor de vitamina C e sua polpa é utilizada em sucos, néctar, geleia, sorvete, bala, iogurte, barra de cereais e bebidas alcoólicas. <sup>79,81,93–96</sup> Na indústria cosmética, tem sido usado em xampus e protetor solar. <sup>79</sup> Na indústria farmacêutica, para produzir comprimidos e cápsulas como fonte natural de vitamina C. <sup>97–99</sup> Atualmente a crescente demanda industrial do camu-camu supera a oferta, o que sugere aumentar áreas de cultivos, uso de tecnologia para evitar perda do fruto após a colheita <sup>80</sup> ou novas formas de obter os derivados que elas produzem naturalmente mediante melhoramento genético, planejamento industrial e técnicas biotecnológicas, com ênfase nas vias metabólicas.



Figura 3 - Fruto de *Myrciaria dubia* "camu-camu," fitoquímicos ativos <sup>84</sup> e seu uso na indústria.
Fonte: Elaborada pelo autor.

Na atualidade, os estudos de sequenciamento genômico, proteômica e transcriptômica de várias espécies de plantas, disponibilizaram informações valiosas da identificação de diferentes vias metabólicas, entre elas as cinco vias metabólicas para a biossíntese de vitamina C em camu-camu. <sup>30,100</sup> Dada a importância industrial de camu-camu, a identificação destas vias metabólicas se tornou crucial para um melhor entendimento das enzimas que estão envolvidas nestes processos biossínteticos. Portanto, o teor excepcional de AsA em camu-camu, comparado com outras espécies de plantas, permite especular a possibilidade de que as enzimas envolvidas na produção de vitamina C de camu-camu podem ter possíveis aplicações biotecnológicas na produção artificial de vitamina C. <sup>30</sup> No entanto, das 5 vias metabólicas a ativação diferencial de um ou mais dependerá de vários fatores (por exemplo, tipo de órgão/tecido, estágio de desenvolvimento, condição fisiológica, fatores ambientais, etc.). Nem todas as vias foram suficientemente estudadas e muitas delas ainda precisam ser demostradas *in vivo*. A via mais estudada até agora é a via D-manose/L-galactose, mesmo assim os estudos moleculares, bioquímicos e sobretudo estruturais são limitados. Portanto, para conseguir futuramente um processo biotecnológico de produção artificial de vitamina C.

caracterizar todas as enzimas que participam da produção de vitamina C de pelo menos uma das vias metabólicas.

Neste sentido, a elucidação das propriedades biofísicas, bioquímicas e estruturais das enzimas envolvidas na biossíntese de vitamina C da via D-manose/L-galactose de *Myrciaria dubia* é crucial para uma melhor compreensão de suas funções. Neste trabalho vamos avaliar parâmetros bioquímicos, biofísicos e estruturais de três enzimas chaves da via D-manose/L-galactose de *Myrciaria dubia*, para nos ajudar a compreender os mecanismos funcionais e regulatórios da via de produção de vitamina C.

#### 1.5 Objetivos

## 1.5.1 Principal

Determinar as propriedades biofísicas, bioquímicas e estruturais da GDP-D-manose-3',5'epimerase, L-galactono-1,4-lactona desidrogenase e da L-galactose desidrogenase da via Dmanose/L-galactose de *Myrciaria dubia* "camu-camu", enzimas chaves na biossíntese de Ácido Ascórbico.

#### 1.5.2 Específicos

- Expressar e purificar as enzimas recombinantes, GDP-D-manose 3',5' epimerase, L-galactono-1,4-lactona desidrogenase e L-galactose desidrogenase.
- Avaliar parâmetros biofísicos das enzimas recombinantes, GDP-D-manose-3',5'epimerase, L-galactono-1,4-lactona desidrogenase e L-galactose desidrogenase usando as técnicas de SEC-MALS e CD.
- > Avaliar parâmetros bioquímicos através da caracterização enzimática das proteínas
- Cristalizar as enzimas recombinantes, GDP-D-manose-3',5'-epimerase, L-galactono-1,4-lactona desidrogenase e L-galactose desidrogenase e determinar as suas estruturas por difração de raios-X.
- Utilizar as informações estruturais obtidas para um melhor entendimento da biossíntese de AsA em plantas.
# **CAPÍTULO II**

# 2 METODOLOGIA

# 2.1 Construção dos Plasmídeos

As sequências do DNA codificante (CDS, do inglês *Coding DNA sequence*) correspondentes as enzimas GDP-D-manose-3',5'-epimerase (*Md*GME), L-galactose desidrogenase (*Md*GDH) e L-galactono-1,4- lactona desidrogenase (*Md*GalDH) de *Myrciaria dubia* com números de acesso do GenBank (**Tabela 1**) foram otimizada para sua expressão em bactérias usando o pacote de software GeneOptimizer <sup>101</sup>, sintetizado por GeneArt® (Invitrogen, Carlsbad, CA, USA) e ligado ao vetor de expressão pET151/D-TOPO. O gene sintético da L-galactose desidrogenase de *Spinacia oleracea* (*So*GDH) com N° de acesso do GenBank (**Tabela 1**) foi ligado ao vetor de expressão plasmidial pET-28a, que foi adquirido da empresa FastBio (São Paulo, Brasil). Portanto, para este trabalho se dispensou o uso das etapas de clonagem dos genes.

Proteína	N° de Acesso GenBank	Vetor de expressão	Resistência ao antibiótico	Massa Molecular Com Cauda	pI	E M <sup>-1</sup> cm <sup>-</sup> 1
<i>Md</i> GME (EC 5.1.3.18)	OK642677.1	pET151/D- TOPO	Ampicilina (100 µg/mL)	46,43	6,03	63,74
<i>Md</i> GDH (EC 1.1.1.316)	OK632632.1	pET151/D- TOPO	Ampicilina (100 µg/mL)	39,30	5,52	31,77
<i>Md</i> GalDH (EC 1.3.2.3)	OK642676.1	pET151/D- TOPO	Ampicilina (100 μg/mL)	61,48	6,13	80,58
SoGDH (EC 1.1.1.316)	AB160990.1	pET-28 <sup>a</sup> (+)	Canamicina (50 µg/mL)	37,55	5,66	29,38

Tabela 1 - Propriedade	s gerais da	as proteínas	estudadas.
------------------------	-------------	--------------	------------

 $\varepsilon$  = coeficiente de extinção molar *pI*=Ponto Isoelétrico

Fonte: Elaborada pelo autor

# 2.2 Expressão e purificação das proteínas recombinantes

Os vetores contendo os genes de interesse (Tabela 1) foram inseridos na cepa de expressão de E. coli Rosseta<sup>TM</sup> (DE3) quimiocompetentes usando a transformação bacteriana por choque térmico. Neste caso, além dos antibióticos dos vetores de expressão, o meio LB-ágar foi suplementado com 34 µg/mL de cloranfenicol (Clo) devido à resistência intrínseca da cepa Rosseta<sup>TM</sup> (DE3). Após o crescimento das células transformadas com os genes de interesse, uma colônia isolada foi inoculada em 10 mL de meio LB contendo os antibióticos (Clo + Amp ou Can) e incubada a 37 °C sob agitação de 90 rpm. Após 16 h, as culturas foram transferidas para um Erlenmeyer contendo 1 L de meio LB com os antibióticos de resistência. As culturas foram incubadas a 37 °C sob agitação a 150 rpm até atingir uma densidade ótica (DO<sub>600nm</sub>) entre 0,6 e 0,8. Atingida a DO, foi induzida a expressão das proteínas recombinantes com 0,3mM de Isopropyl-β-D-thiogalactopyranoside (IPTG) por 16 h a 20 °C sob agitação de 150 rpm. Após as 16 h de incubação, as células foram coletadas por centrifugação a 4000 rpm x 45 min a 4 °C. Em seguida, as células foram ressuspendidas em 40 mL do tampão A (20 mM Tris/HCl pH 7,5; 150 mM NaCl; 10% glicerol) e submetidas a sonicação com pulsos de ultrassom de 14 ciclos de 30 segundos e intervalos de 59 segundos por cada ciclo de sonicação. A fração solúvel foi separada da insolúvel por centrifugação a 13 000 rpm por 45 min a 4 °C.

A purificação foi realizada em duas etapas, a primeira etapa de purificação foi realizada por cromatografia de afinidade por metal imobilizado (IMAC), o qual é facilitada pela presença de uma cauda de 6xHis na proteína recombinante. Utilizou-se uma coluna contendo 3 mL de Ni-NTA Agarose (Quiagen®) equilibrada com 5 volumes de coluna do tampão A. Após a imobilização das proteínas recombinantes na resina, foram feitas lavagens com 5x o volume do tampão A e 5 passos de eluição seriada, onde o tampão A continha concentrações crescentes de imidazol (25 mM - 50 mL, 50 mM - 25 mL, 100 mM - 15 mL, 250 mM - 10 mL, 500 mM - 5 mL), as eluições foram coletadas e analisadas por eletroforese em gel de poliacrilamida 12% sob condições desnaturantes (SDS-PAGE).

A segunda etapa de purificação foi realizada por cromatografia de Exclusão Molecular (*SEC*). As eluições com major grau de pureza, provenientes da cromatografia de afinidade, foram agrupadas e concentradas até 2 mL através do concentrador Amicon® Ultra (limite de peso molecular 30 kDa) (Merck Millipore, Darmstadt, Alemanha), e injetadas numa coluna Superdex 200 XK16/60 (GE Healthcare) pré-equilibrada com tampão A, acoplada num sistema *ÄKTA purifier* (GE Healthcare). A cromatografia foi realizada com fluxo de 1

mL/min, com uma eluição isocrática de 1,5X do volume total da coluna, monitoradas a 280<sub>nm</sub> e os picos de eluição foram coletados e analisados por SDS-PAGE 12%.

# 2.3 Caracterização biofísica

# 2.3.1 Determinação do estado oligomérico por cromatografia de exclusão molecular acoplada ao espalhamento e luz a múltiplos ângulos (SEC-MALS)

Para a determinação da massa molecular e do estado oligomérico das proteínas recombinantes em estudo, foi utilizado a técnica de Cromatografia de exclusão molecular acoplada ao espalhamento de luz de múltiplos ângulos (SEC-MALS, Multiple-Angle Light Scattering). A massa molecular é calculada usando uma curva de calibração com uma proteína com peso molecular conhecido. A técnica estima o volume das partículas em solução à medida que elas espalham a luz em várias direções. <sup>102</sup> O cálculo feito utilizando está técnica permite obter a massa molecular de qualquer tipo de proteína, inclusive agregados intrinsecamente desordenados. <sup>103</sup>

As proteínas eluidas da SEC, foram concentradas a >2mg/mL, centrifugadas a 10000rpm por 10 min a 4 °C. Foram injetados 100 µL de cada amostra numa coluna de exclusão molecular *Superdex 200 Increase 10/300 GL (GE Healthcare)*, equilibrada com tampão Hepes 25 mM pH 7.8, NaCl 300 mM e 5 MgCl<sub>2</sub> 5mM. A coluna cromatográfica foi acoplada a um sistema de Cromatografia Líquida de Alta Eficiência (HPLC) *Waters 600 (Wyatt Technology Corporation)*, um módulo de espalhamento estático de luz a múltiplos ângulos miniDAWN® TREOS®- MALS (*Wyatt Technologies*), um detector WyattQELS® para medidas de espalhamento dinâmico de luz (DLS, Dynamic Light Scattering) e um detector de índice de refração, refratômetro, Optilab® T-rEX (*Wyatt Technology Corporation*). A corrida cromatográfica foi realizada com fluxo de 0,5 mL/min, em uma temperatura de 20 °C, as análises do MALS/DLS são monitoradas em tempo real, determinando de forma direta e absoluta a Massa molecular, índice de refração da solução (n) e mudança do índice de refração de vido à concentração da amostra. <sup>104</sup> Os dados foram analisados no programa *ASTRA 7.0.1* (*Wyatt Technology*).

# 2.3.2 Análise do enovelamento e estabilidade térmica com espectroscopia de dicroísmo circular (CD)

Com a finalidade de investigar o comportamento em solução das proteínas de estudo, utilizou-se a espectroscopia de Dicroísmo Circular (CD), técnica baseada na interação diferenciada de moléculas assimétricas com luz circularmente polarizada no sentido horário (RCP) e anti-horário (LCP), num espectro de 190 nm e 270nm. <sup>105–107</sup>

Os espectros de dicroísmo circular (CD) foram obtidos mediante um espectropolarímetro J-815 (JASCO, Corporation, Tóquio-Japão). O UV distante (Far/UV) da amostra foi monitorado no intervalo de comprimento de onda entre 190 até 280 nm, resolução de 0,2 nm, numa cubeta de quartzo (caminho óptico de 0,1cm), acoplado ao sistema de variação térmica *Peltier* com parâmetros de 50 nm/min. As concentrações das proteínas utilizadas foram de ~0,250 mg/mL, em tampão fosfato de sódio-Na (20 mM, pH 7,5 -7,8), contendo NaCl 100 mM e glicerol a 5%. O ponto médio da transição (T<sub>m</sub>)entre o estado enovelado e desenovelado das proteínas, foi determinado submetendo as amostras a um gradiente de temperatura de 4 a 94°C (como máximo), medidos em intervalos de 2 °C. A curva de *melting* foi estipulada a partir da mudança do sinal de CD em função da temperatura usando o comprimento de onda específico (222 nm, mínimo correspondente às hélices-α).

Todos os espectros foram registrados usando o software *Spectra Manager* (JASCO), a análise dos espectros foi realizado mediante o programa CDTools <sup>108</sup>, onde os espectros originais foram subtraídos das contribuições do solvente, tratados pelo algoritmo de Transformada Rápida de Fourier (FFT- Fast Fourier Transform). Os gráficos foram plotados no software OriginLab <sup>109</sup>, versão 9.9.0.225.

#### 2.3.3 Análise de clivagem de *Md*GME

A finalidade deste experimento foi investigar a possível clivagem de *Md*GME, devido a presença de duas bandas muito próximas que sempre eram detectadas nos géis de SDS-PAGE, após as purificações. A análise foi feita com o espectrômetro de massas modelo micrOTOF-QII – Bruker Daltonics®. Sendo as amostras preparadas da seguinte forma: As bandas do SDS-PAGE correspondentes a *Md*GME foram cortadas e fragmentadas em tamanhos de aproximadamente 1mm. As amostras foram descoradas completamente com uma solução de metanol 50% e ácido acético 2,5 %. Após o descoramento, adicionou-se uma solução de 100 mM de bicarbonato de amônio e 50% de acetonitrila (ACN), em seguida, desidratou-se a

amostra com 200  $\mu$ L ACN 100% duas vezes, a solução foi removida por evaporação, e foi adicionado de 30 a 50  $\mu$ L de solução 10 mM de DTT, e incubado por 30 min a temperatura ambiente, a solução é então removido e adicionado de 30 a 50  $\mu$ L de 50 mM de IAA (Iodoacetamida) e incubado novamente por 30 min a temperatura ambiente, ambas as soluções (DTT e IAA) foram homogeneizados em 1mL de bicarbonato de amônio 100 mM, a remoção foi realizada com 100  $\mu$ L de bicarbonato de amônio 100 mM. Após este passo, realizou-se novamente a desidratação da amostra e remoção da ACN, em seguida as amostras foram reidratadas com 50  $\mu$ L de uma solução de 50 mM de bicarbonato de amônio gelado suplementado com 20  $\mu$ g de tripsina [20ng/  $\mu$ L), e incubadas 30 min em gelo, o excesso da solução foi removido e logo em seguida adicionou-se 50 mM de bicarbonato de amônio até cobrir a amostra, as amostras foram incubadas por 16 h a 37 °C, esta etapa permite obter os peptídeos a partir das proteínas clivadas pela tripsina, que posteriormente serão identificadas.

Após incubação, foi adicionado uma solução 5% de ácido fórmico e incubado por 10 min a temperatura ambiente, para a atividade da tripsina ser interrompida e as amostras foram centrifugadas e o sobrenadante transferido para um novo microtubo. Em seguida foi adicionada uma solução de ácido fórmico ao 5% e ACN 50% sob o pellet do passo anterior e incubadas por 10 min a temperatura ambiente, seguido por um "spin" de 30 seg. O sobrenadante obtido foi misturado com o sobrenadante do passo anterior. Feito isto, as amostras foram submetidas a um evaporador (*speed vac*) até atingir um volume ~50  $\mu$ L e congeladas a -20 °C até a realização da análise. A amostras dos peptídeos foram aplicadas no espectrômetro de massa micrOTOF-QII – Bruker Daltonics®, e os resultado foram analisados utilizando o software *MASCOT* (*http://www.matrixscience.com/search\_form\_select.html*).

#### 2.3.4 Ensaios de cristalização, coleta de dados e resolução das estruturas

A cristalografia é uma das principais técnicas utilizadas para a determinação do modelo atômico experimental de macromoléculas em alta resolução em um arranjo cristalino, usando raios-X como ferramenta de medição. <sup>110</sup> Precisamente em cristalografia de proteínas, a formação de cristais proteicos depende de uma ampla gama de condições incluindo as concentrações dos componentes (proteína-precipitante), a solubilidade (proteína), a presença de aditivos (sais, ligantes e metais), a temperatura de incubação etc. <sup>111-112</sup>

Os ensaios de cristalização das proteínas recombinantes foram realizados usando a técnica de difusão de vapor por gota sentada automatizada, através do robô de cristalização *Crystal Gryphon (Art Robbins Instruments*®), em placas *Intelli-plate 96-low profile (Art Robbins* 

Instruments), e com os kits de cristalização comerciais (Morpheus MD®, BCS Screen MD®, Index HR®, JCSG MD® e SG1 MD®). A proporção da solução de cristalização e proteína foi de 1:1 com os volumes de 0.2/0.2 µL, com 50 µL de solução no reservatório. A concentração da proteína variou de 2 mg/mL até 25 mg/mL, sendo testadas condições apo e com os ligantes respectivos de cada proteína (NAD<sup>+</sup>, FAD, Citocromo C, L-galactose, L-galactono 1,4lactone e GDP-D-mannose), as placas foram seladas com um filme adesivo e armazenadas no hotel de placas Rock Imager 1000 (Formulatrix), o equipamento possui um sistema de imagem automatizado incorporado com luz UV/visível para diferenciação de cristais de sal e proteína, além de ter uma temperatura regulada (20 °C). Os experimentos são mantidos no sistema com uma rotina de visualização programada seguindo a série de Fibonacci por 55 dias (0, 1, 2, 3, 5, 8, 13, 21, 34 e 55). Quando obtidos cristais bons (grandes e isoladas) os mesmos foram resfriados rapidamente e mantidas em nitrogênio líquido. Enquanto as condições cujas amostras resultaram em cristais proteicos pequenos, aglomerados e poli cristais tiveram seus parâmetros submetidos a um refinamento das condições originais na tentativa de melhorar os cristais. Para este processo se usaram placas VDXm Plate de 24 poços (Hampton Research) e a técnica de difusão de vapor por gota sentada e gota pendurada. Com as proporções de 1:1 (1  $\mu$ L/1  $\mu$ L) de solução de cristalização e proteína, 250 a 500  $\mu$ L de solução no reservatório, diferentes concentrações de proteína >10 mg/mL, variações do pH e porcentagem de precipitante nas soluções foram testados.

Os cristais obtidos foram retirados da gota de cristalização cuidadosamente com um *MicroLoop (MiTeGen)* de tamanho compatível com o tamanho do cristal. Antes do resfriamento, os cristais sem crio-protetor na condição de cristalização, foram submetidos a uma solução crio-protetora formada por 20 % de glicerol, PEG 200 ou PEG 400 e 80% da solução de cristalização respectiva.

Os dados de difração de raios-X foram coletados a 100 K no síncrotron Sirius (LNLS-CNPEM, Campinas) usando a linha de luz Macaná abrigando um detector PILATUS 2M para *Md*GalDH, *Md*GDH e *So*GDH, ou no síncrotron Diamond Light Source (Didcot, Reino Unido) usando a linha de luz I03 com um detector Eiger2 XE para *Md*GME-NAD<sup>+</sup> e *So*GDH-NAD<sup>+.</sup> Os dados foram indexados, integrados e escalonados usando Xia2. <sup>113</sup> Cada estrutura foi processada de forma distinta, começando com *So*GDH-apo. As coordenadas para *Os*GDH (PDB: 7EZI e 7EZL) ainda não estavam disponíveis no momento da solução da estrutura e, portanto, tentativas de faseamento, usando substituição molecular convencional, foram feitas empregando Peracina redutase (PDB:3V0T) como modelo de busca. Estes não tiveram sucesso e, em vez disso, adotamos o "*pipeline*" do *MoRDa*. <sup>114</sup> O *MoRDa* usa o banco de dados (2G) derivado do PDB que contém uma descrição compacta de cadeias de proteínas não redundantes, domínios, homo e hetero-oligômeros. Os modelos de domínio préprocessados para excluir os loops flexíveis e otimizar as chances de sucesso da substituição molecular. A solução da estrutura é automatizada usando o programa *MOLREP*. <sup>115</sup> No caso de *So*GDH, várias estruturas foram usadas no processo, incluindo 1YNP (29,3% de identidade), 3UYI (30,6% de identidade), 6HG6 (26,1% de identidade), etc. O melhor modelo (baseado em Z-Score e valores finais de R<sub>work</sub> e R<sub>free</sub>) foi submetido à construção automatizada de modelos com *ARP/wARP* <sup>116</sup> para obter um modelo final com a sequência correta. Posteriormente, *So*GDH-NAD<sup>+</sup> foi resolvido por substituição molecular com a estrutura *So*GDH-apo previamente refinada como modelo de busca usando Phaser. <sup>117</sup> Rodadas alternadas de refinamento e reconstrução de modelo foram realizadas usando Phenix <sup>118</sup> e Coot <sup>119</sup> produzindo assim os modelos finais. As figuras foram geradas com PyMol v2.05 (Schrödinger, LLC). <sup>120</sup> Os resíduos envolvidos no sítio de ligação NAD<sup>+</sup> foram determinados com o Discovery Studio Visualizer V21.1.0 (BIOVIA, Dassault Systèmes). <sup>121</sup>

Para *Md*GalDH, as coordenadas da subunidade GalDH extraída do complexo mitocondrial I de *Brassica oleracea-Bo*GalDH (PDB:7A24, identidade 83,2%) foram testadas para faseamento, usando substituição molecular convencional sem resultados promissores. Por esse motivo, foi realizada a predição da estrutura de *Md*GalDH usando AlphaFold2, o modelo foi usado para substituição molecular usando Phaser <sup>117</sup> do pacote Phenix. <sup>118</sup> A mesma abordagem foi usada para substituição molecular de *Md*GME. Como mencionado anteriormente, ciclos de refinamento e construção do modelo foram alternados usando Phenix <sup>118</sup> e Coot, <sup>119</sup> produzindo os modelos finais.

# 2.4 Caracterização bioquímica

#### 2.4.1 Ensaios cinéticos de GDHs

A cinética enzimática, os ensaios de inibição e a influência do pH foram realizados em um espectrofotômetro de microplacas SpectraMax Plus 384 (Molecular Devices, Sunnyvale, Califórnia, EUA). A conversão de NAD<sup>+</sup> em NADH ( $\mathcal{E}=2650 \text{ M}^{-1}.\text{cm}^{-1}$ ) foi monitorada em um microplaca transparente de UV de 96 poços, registrando a absorbância a 340 nm a cada 6 seg por 5 min a 26 °C. As velocidades iniciais da enzima foram estimadas como as taxas lineares máximas de aumento de absorbância em 340 nm. Os valores de mAU foram convertidos para concentração molar usando a equação de Beer-Lambert  $\mathbf{A} = \boldsymbol{\epsilon} \cdot \boldsymbol{\ell} \cdot \mathbf{c}$ , onde: A

é a absorbância,  $\varepsilon$  é o coeficiente de extinção molar de NADH a 340 nm,  $\ell$  é o comprimento do caminho óptico em cm e *c* é a concentração molar de NADH. Em seguida, esses valores foram convertidos em unidades de taxa de reação (concentração/tempo).

A determinação do pH ideal das enzimas foi determinado na faixa de 4 a 11. A atividade catalítica foi medida como mencionado acima, em 200  $\mu$ L de mistura de reação em um tampão universal (uma mistura de três ácidos tri-próticos: ácido cítrico, ácido bórico e ácido fosfórico) <sup>122</sup> contendo 200  $\mu$ M de NAD<sup>+</sup>, L-galactose (L-gal, 1mM para *So*GDH e 2mM para *Md*GDH) e a enzima recombinante purificada foi adicionada a uma concentração final de 100nM.

A análise de cinética enzimática *So*GDH e *Md*GDH recombinantes foram testados em uma mistura de reação de 200  $\mu$ L 100 mM de Tris-HCl pH 7.0, 200  $\mu$ M de NAD<sup>+</sup>, L-gal (1 mM para *So*GDH e 2 mM para *Md*GDH) e enzima purificada a 100 nM. Os parâmetros cinéticos enzimáticos (K<sub>m</sub>, V<sub>max</sub>, V<sub>max</sub>/K<sub>m</sub> e k<sub>cat</sub>) foram determinados através ajuste de regressão não linear dos dados experimentais usando o software Prism - GraphPad v9.00.

Para avaliar o efeito inibitório do ácido L-ascórbico na atividade da GDH, 200 µL de reação contendo 100 mM de Tris-HCl pH 7.0, 200 µM de NAD<sup>+</sup>, L-gal (1 mM para *So*GDH e 2 mM para *Md*GDH) e 100 nM de enzima purificada na presença de várias concentrações de AsA (0.75, 1.5, 3.0 e 4.5 mM) foram monitorados como descrito acima. Da mesma forma, a inibição também foi monitorada usando um tampão mais concentrado (300 mM Tris-HCl pH 7.0). A porcentagem de inibição foi calculada usando a seguinte equação: *Inibição* (%) =  $\frac{(\alpha-1)}{(\alpha+\beta)}x100$ , onde:  $\alpha \in (1+[I])/Ki$ ,  $\beta$  is [S]/K<sub>m</sub>, [I] é a concentração do inibidor e [S] é a concentração do substrato.

#### 2.4.2 Ensaios cinéticos MdGalDH

A cinética enzimática *Md*GalDH foi realizado de forma similar a GDH. Os ensaios foram realizados seguindo a redução de citocromo C ( $\mathcal{E}=21 \text{ mM}^{-1}.\text{cm}^{-1}$ ) em um espectrofotômetro de microplacas SpectraMax Plus 384 e monitorada em uma microplaca transparente de 96 poços, as absorbâncias foram registradas a 550 nm a cada 6 seg por 10 min a 26 °C. As velocidades iniciais da enzima foram estimadas como as taxas lineares máximas de aumento de absorbância em 550 nm. Os valores de mAU foram convertidos para concentração molar usando a equação de Beer-Lambert A= $\epsilon$ . 1. c, onde: "A" é a absorbância, " $\mathcal{E}$ " é o coeficiente de extinção molar do Citocromo C (Cc) reduzido a 550 nm, "1" é o comprimento do caminho

óptico em cm e "c" é a concentração molar do citocromo C reduzido. Em seguida, esses valores foram convertidos em unidades de taxa de reação (concentração/tempo).

A determinação do pH ideal foi determinado na faixa de 4 a 11. A atividade catalítica foi medida como mencionado acima, utilizando 200  $\mu$ L de reação em um tampão universal (uma mistura de três ácidos tri-próticos: ácido cítrico, ácido bórico e ácido fosfórico)<sup>77</sup> contendo 50  $\mu$ M de Citocromo C, 0,3-3 mM de L-galactono-1,4-lactona (L-galL) e 100 mM de L-gulono-1,4-lactona (L-gulL), a enzima recombinante purificada foi adicionada a uma concentração final de 5 e 50nM respectivamente.

A análise de cinética enzimática da proteína *Md*GalDH recombinante foi testado em 200  $\mu$ L de reação contendo 100 mM de Tris-HCl pH 8.0, 50  $\mu$ M de Citocromo C, L-galL (0.3 mM) e enzima purificada a 5 nM. Os parâmetros cinéticos enzimáticos (K<sub>m</sub>, V<sub>max</sub>, V<sub>max</sub>/K<sub>m</sub> e k<sub>cat</sub>) foram determinados através do ajuste de regressão não linear dos dados experimentais usando o software Prism- GraphPad v9.00.

Para avaliar o efeito inibitório do L-galactono-1,4-lactona (L-galL) na atividade da proteína *Md*GalDH, 200 µL da reação contendo 100 mM de Tris-HCl pH 8.0, 50 µM de Citocromo C, 3 mM L-galL e 50 nM de enzima purificada, foram monitoradas como descrito anteriormente. A porcentagem de inibição foi calculada usando a seguinte equação: *Inibição* (%) =  $\frac{(\alpha-1)}{(\alpha+\beta)}x100$ , onde:  $\alpha$  é (1+[i])/Ki,  $\beta$  is [S]/K<sub>m</sub>, [i] é a concentração do inibidor e [S] é a concentração do substrato.

#### 2.5 Análises de sequências homólogas

Cada uma das sequências das proteínas (*Md*GME, *Md*GDH, *So*GDH e *Md*GalDH) foram analisadas com sequências homólogas com identidade na faixa entre 25 e 90%. O alinhamento múltiplo de sequências foi realizado usando o software CLUSTAL W.<sup>123</sup> Por exemplo: No caso das GDHs, a história evolutiva de 13 subfamílias de AKRs anotadas do banco de dados PKR da Universidade da Pensilvânia e 5 representantes de plantas L-GHDs foi inferida usando o método de máxima verossimilhança.<sup>124</sup> Uma árvore construída com as sequências consenso "bootstrap" foi obtida a partir de 1000 réplicas, <sup>125</sup> com ramos da subfamília colapsados sempre que possível e correspondentes a partições reproduzidas em menos de 30% de réplicas de bootstrap também foram recolhidos. As análises filogenéticas foram realizadas no MEGA11.<sup>126</sup>

No caso de *Md*GME e *Md*GalDH o objetivo foi identificar os resíduos de aminoácidos conservados do sítio ativo com seus homólogos (GMEs e GalDHs).

# 2.6 Previsão da estrutura da proteína

Um caso especial foi levado a cabo com a proteína *Md*GDH. Devido a que os cristais que obtivemos não eram de qualidade o suficiente para fazer experimentos de difração de raios-X, realizamos a previsão de sua estrutura com o ColabFold, uma extensão online do software AlphaFold2. <sup>127-128</sup> A busca MMseqs2 <sup>129</sup> selecionamos os parâmetros de entrada, ((uniref+environmental) msa\_mode, automatic model\_type, unpaired+paired pair\_mode and 3 rum\_recycles). 150 sequências foram selecionadas como templates para construção dos modelos, obtendo-se 5 modelos finais. A qualidade das estruturas proteicas tridimensionais modeladas foi avaliada por parâmetros IDDT <sup>127-128</sup>, estereoquímica por Procheck <sup>130</sup> e Verify3D. <sup>131</sup>

# **CAPÍTULO III**

#### **3 RESULTADOS E DISCUSSÕES**

Para uma melhor compreensão, interpretação e discussão dos dados, estes foram divididos em 3 subcapítulos, cada subcapítulo apresentará os resultados, discussões e conclusões de cada uma das enzimas de estudo. Ressalta-se que grande parte dos resultados descritos no **subcapítulo 3.2** foram recentemente publicados em um artigo científico. <sup>132</sup>

# 3.1 Subcapítulo: GDP-mannose -3', 5'-epimerase (GME)

#### 3.1.1 Purificação e determinação do estado oligomérico de MdGME

Conforme descrito na metodologia, a expressão heteróloga e purificação de camu-camu *Md*GME foi realizada através de duas etapas. A primeira, constituída por uma cromatografia de afinidade ao níquel, essencial para isolar a proteína de interesse de todas as outras expressas da própria maquinaria da bactéria *E. coli*. As bandas observadas no gel de SDS-PAGE (**Figura 4A**) mostram uma proteína solúvel, com uma banda próxima a 45 kDa (massa esperada de 46,43 kDa). Este resultado mostra que *Md*GME purifica com uma baixa quantidade de contaminantes como observado nas frações de eluição E2 e E3 (**Figura 4A**).

Na segunda etapa realizou-se uma cromatografia de exclusão molecular, onde as eluições que apresentaram uma banda única na etapa 1, foram concentradas e injetadas numa coluna HiLoad Superdex 200 16/60. O cromatograma da purificação (**Figura 4B**) mostra um pico bem definido e, a pureza foi analisada por gel de SDS-PAGE 12%. No entanto, observamos duas bandas com tamanhos próximos à banda do marcador de 45 kDa (**Figura 4**), o que indicaria uma possível degradação ou clivagem da proteína. Para identificar a presença de clivagem, as bandas foram cortadas e tratadas com tripsina e analisadas por espectrometria de massas (**Figura 5**). Para ambas as bandas, após um mapeamento de peptídeo-massa pelo programa *Mascot*, as sequências apresentaram uma cobertura de 41% com uma proteína prevista de 42,5 kDa de *Oriza sativa* anotada como GME. Portanto, as duas bandas correspondem a versões diferentes de *Md*GME. A massa molecular de *Md*GME é de 46,43 kDa, portanto, a banda superior observada no gel representa a proteína GME com sua massa

total. A segunda banda, identificada como GME com uma massa de 42,5 kDa, estaria relacionada com um fenômeno de clivagem que já foi relatado para esta proteína de *Arabidopsis thaliana*. <sup>43</sup> Esta clivagem acontece no extremo C-terminal da proteína perdendo aproximadamente 30 aminoácidos (~3 kDa), coincidindo com a massa molecular calculada pelo espectrômetro de massas (**Figura 5**). No caso de *Md*GME não foi possível identificar a posição da clivagem, porém parece provável que seja análoga àquilo observada em *Arabidopsis thaliana* com a perda da região C-terminal.



Figura 4 - Expressão e purificação de MdGME. (A) SDS-PAGE da purificação por cromatografia de Afinidade de MdGME M:Marcador de massa molecular, AI: antes de induzir com IPTG, DI: depois de induzir com IPTG, P: pellet-fração Insolúvel, S: fração solúvel, NL: proteínas que não interagem com o níquel imobilizado na resina, L1: lavagem com tampão, E1-E5: eluições com 25,50,100,250 e 500 mM de imidazol. B. Cromatografia de Exclusão de tamanho (SEC) de MdGME.

Fonte: Elaborada pelo autor.



Figura 5 – Espectrometria de massa de MdGME. O "output" da análise, usando o Mascot Research. Em vermelho estão ressaltados os peptídeos identificados a partir da banda indicada no gel na direita da figura.

Fonte: Elaborada pelo autor.

O principal objetivo deste trabalho de pesquisa é estudar as propriedades bioquimicas (relacionadas à atividade da proteína) e, biofísicas (oligomerização, estabilidade termica e estrutura tridimencional) da proteína de *Md*GME. Portanto, definir a homogeneidade e estado oligomerico de nossa proteína é importante para dar continuidade as proximas etapas de caracterização. Para a determinação do estado oligomerico da enzima recombinante de *Md*GME usamos a cromatografia de exclusão por tamanho acoplada aos perfis de espalhamento de luz multi-ângulo (SEC-MALS) como tecnica absoluta para determinar a massa molecular de proteína em solução, assim como a homogeneidade da amostra. O perfil de SEC-MALS mostrou um pico homogêneo, o qual corresponde a uma molecula de 94,22 kDa. Esta massa é corente com um dimero de GME, uma vez que a massa molecular teórica do monomero é 46,43 kDa (**Figura 6**). Até o momento estudos relacionados ao estado oligomerico de homólogos a *Md*GME, foram relatados em *Arabidopsis thaliana* (*At*GME). A proteína *At*GME purificada apresentou uma massa de 84 kDa, composto de duas subunidades de 42,8 kDa. <sup>43-44</sup> Desta forma, nossos resultados confirman o estado oligomerico de GME.



Figura 6 - Cromatograma obtido através do SEC-MALS de *Md*GME. Em vermelho temos as curvas correspondentes à mudança no índice de refração diferencial normalizado. Em azul a intensidade de luz espalhada normalizada a 90°. Em preto a massa molecular calculada através do pico correspondente. A MW determinado corresponde ao valor esperado para um homodímero de *Md*GME.

Fonte: Elaborada pelo autor.

#### 3.1.2 Enovelamento e estabilidade térmica de *Md*GME

Avaliamos o estado de enovelamento da proteína mediante o monitoramento do seu espectro de CD, a cada 2°C, em função do aumento de temperatura usando um gradiente de 10 a 70°C. A **Figura 7A**, mostra os espectros de CD de *Md*GME em temperaturas onde houve

mudanças do seu sinal. A 20°C, podemos observar mínimos negativos em 222 e 208 nm, predominância do sinal referente às contribuições de hélices a. Os espectros após 56°C mostram uma diferença significativa, que está relacionado com a perda de elementos de estrutura secundária em MdGME. A curva de desnaturação térmica, onde o sinal do espectro de CD foi monitorado a 222nm, mostra a temperatura de transição ou chamada também de temperatura de melting (T<sub>m</sub>), obtida (59.4°C  $\pm$  0,4) após um ajuste de *Boltzman* no sinal detectada (Figura 7B), o processo é irreversível e a proteína precipita a altas temperaturas. Em contraste, estudos de T<sub>m</sub> em outras proteínas GME de plantas como Oryza sativa (OsGME) e AtGME mostraram um valor de T<sub>m</sub> de 25°C e 36°C, <sup>56,59</sup> respectivamente. No entanto, Methylacidiphilum fumariolicum (MtGME) apresentou um T<sub>m</sub> de 57 °C, <sup>59</sup> uma temperatura relevante para a indústria de produção de açucares de difosfato de nucleosídeos (NDP). Embora a medidas da T<sub>m</sub> tenha sido determinada usando técnicas diferentes, a variação da T<sub>m</sub> não seria muito grande. Nesse sentido, tanto MtGME, como MdGME poderiam ser considerados enzimas com uma temperatura promissora para exploração industrial de produção de L-acúcares ou derivados (Vitamina C).



**Figura 7** - Espectro de CD de *Md*GME. (A) Os espectros de CD mais representativos obtidos entre 10°C a 70°C. (B) Curva de desnaturação térmica de *Md*GME,  $T_m 59.4^\circ$ C ± 0,4 (caixa vermelha).

Fonte: Elaborada pelo autor.

# 3.1.3 Descrição geral da estrutura cristalográfica de MdGME

Até a data de escrita deste trabalho de dissertação, as únicas estruturas cristalográficas reportadas na literatura para GME foram de *Arabidopsis thaliana*. <sup>55</sup>

Neste trabalho de pesquisa abordamos diversas tentativas de cristalização (com cofator, sem cofator, com substrato, sem substrato e cofator/substrato), das quais as condições mais favoráveis que resultaram em cristais com qualidade de difração de raios-X estão ilustradas na **Figura 8**. Dados de difração de qualidade foram coletados para *Md*GME ligada a seu cofator NAD<sup>+</sup> (*Md*GME-NAD<sup>+</sup>) e para *Md*GME na presença de NAD<sup>+</sup> e seu substrato GDP- $\alpha$ -D-Manose (*Md*GME-NAD<sup>+</sup>-GDD/GDC). A nomenclatura usada aqui será justificada mais adiante. As estruturas foram refinadas até resoluções de 2,58 e 1,25 Å, respectivamente e as estruturas finais obtidas foram de boa qualidade. A estrutura *Md*GME-NAD<sup>+</sup> apresentou valores de R<sub>work</sub>= 21,16% e R<sub>free</sub>= 24.51%, com 96.93% de aminoácidos nas regiões mais favorecidas no Ramachandran Plot. Por outro lado, *Md*GME-NAD<sup>+</sup>-GDD/GDC, apresentou valores de R<sub>work</sub>= 14.75% e R<sub>free</sub>= 16.51%, com 97.76% de aminoácidos nas regiões mais favorecidas no Ramachandran Plot. Um resumo dos parâmetros da coleta e estatística de refinamento são mostrados na **Tabela 2**.



Figura 8 - Cristais da proteína de *Md*GME. (A)Exemplo dos cristais obtidos com qualidades de difração por raios X para a condição A10 (0,2 M Sulfato de Litio, 0,1 M Tris 8,5 30% w/v PEG 4000) do kit de cristalização SG1. (B e D) Exemplo dos cristais obtidos com qualidades de difração por raios X com modificações da condição original (A10-SG1) para obtenção de um complexo só com NAD<sup>+</sup> (A10-SG1+5mM NAD<sup>+</sup>). (C)Exemplo dos cristais obtidos com qualidades de difração por raios X para a condição E1 (0,2 M Sulfato de Amonio, 0,1 M Bis-Tris 5,5) do kit de cristalização SG1. Todos foram cristalizados com [~3 a 7mg] de proteína (sem cauda de Histidina).

Fonte: Elaborada pelo autor.

	MJCME NAD+	MdGME-NAD+-
	MaGME-NAD	GDD/GDC
Fonte de X-ray	Sirius MANACA	Diamond I03
Detetor	Pilatus 2M	Eiger2 XE
Parâmetros da cela: a, b, c (Å)	87.85 80.04 107.66	89.80, 60.84, 133.99,
Parâmetros da cela: $\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 95.25, 90,00	90.00, 105.86, 90.00
Grupo espacial	$P2_1$	<i>C</i> 2
Resolução (Å)	107.20 - 2.58 (2.67 - 2.58)	44.81-1.25 (1.295 - 1.25)
λ (Å)	1.33	0.9795
Multiplicidade	6,3 (6,0)	6.4 (3.6)
$R_{\rm pim}$ (all I+ & I-) (%)	12,1 (60,0)	3.1 (44.2)
CC(1/2)	0.982 (0.524)	0.999(0.648)
Completeza (%)	100.0 (99.9)	96.4 (67.7)
Reflexões	294769(27532)	1179222 (23008)
Reflexões únicas	47036 (4579)	184640 (6343)
$< I/\sigma(I) >$	5.6 (1.3)	12.1 (1.4)
R (%) *	21,16	14.75
$R_{\rm free}$ (%) *	24.51	16.51
Número de átomos: proteína	1296	5755
Número de átomos: Água	299	1089
Número de átomos: ligante	176	254
B (Å <sup>2</sup> )	38,14	9,29
Erro de coordenada (basado no-ML) (Å)	0.34	0,10
Erro de fase (°)	26.43	16.12
Ramachandran favorecida (%)	96.93	97.76
Ramachandran permitida (%)	2.76	1.96
Pontuação de confrontos de todos os átomos	6.02	3.07
Comprimento de ligações (RMSD) (Å)*	0.004	0.013
Ângulo de ligações (RMSD) (°) *	0.65	1.37

Tabela 2- Coleta de dados de raios-X e estatísticas de refinamento de estrutura de MdGME

Fonte: Elaborada pelo autor

# 3.1.4 Estrutura de *Md*GME-NAD<sup>+</sup>-GDD/GDC

A primeira estrutura de *Md*GME resolvida neste trabalho corresponde ao cristal gerado na presença de NAD<sup>+</sup> e seu substrato, GDP- $\alpha$ -D-manose (GDD). A determinação da estrutura revelou densidade no sítio ativo que correspondia a uma mistura de substrato (GDD) e produto, GDP- $\beta$ -L-galactose (GDC) uma condição incomum e que será discutida mais

adiante. Por este motivo será usado o nome *Md*GME-NAD<sup>+</sup>-GDD/GDC durante o restante do texto. Esta estrutura cristalográfica apresentou duas moléculas na unidade assimétrica. Após análise por PISA <sup>133-134</sup> e comparação com seu homólogo *At*GME, conclui-se que se trata de um dímero fisiológico, coerente com os dados de SEC-MALS. A estrutura de *Md*GME-NAD<sup>+</sup>-GDD/GDC foi resolvida por substituição molecular usando a estrutura de *At*GME <sup>55</sup> como modelo de busca através do programa *Phaser* no pacote *Phenix*. <sup>117-118</sup>

A proteína *Md*GME faz parte da superfamília desidrogenase/redutase de cadeia curta (SDR) estendido, cujos membros se caracterizam por apresentar um motivo G<sub>XX</sub>G<sub>XX</sub>G para a ligação do NAD<sup>+</sup> no N-terminal de seu domínio *Rossmann-fold* e um motivo catalítico Y<sub>XXX</sub>K no C-terminal do mesmo domínio.<sup>135</sup> Um estudo baseado em *Hidden Markov models* (HMM) classifica as SDR em 464 famílias, <sup>136</sup> das quais estão presentes nos reinos archaea (50), bactéria (374) e eucariota (236). Um grupo de proteínas importantes dentro destas famílias são as enzimas metabolizadoras de açúcar.<sup>136</sup> Uma revisão baseada nestas enzimas, epimerase de carboidratos (CEP, EC 5.1.3), revela uma nova classificação baseada na estrutura e mecanismo das especificidades de CEP.<sup>137</sup> As proteínas *Md*GME estão agrupadas na família CEP 1 (epimerase/desidratase dependente de NAD<sup>+</sup>), juntamente com UDP-glicose 4-epimerase, CDP-paratose 2-epimerase, UDP-glucosamina 4-epimerase, entre outras.<sup>137</sup> A família CEP1 é de longe a maior família. Todos os membros desta família utilizam um mecanismo intermediário ceto transitório, o qual se caracteriza pelo uso de NAD<sup>+</sup> ou NADP<sup>+</sup> como cofator, para subtrair o hidreto da posição de epimerização, além de uma tétrade Asn-Ser-Tyr-Lys conservada em seu sítio ativo.<sup>137-138</sup>

A estrutura de *Md*GME é dividida em dois domínios, o domínio *Rossmann-fold* de ligação a NAD<sup>+</sup> o qual possui 7 fitas  $\beta$  paralelas flanqueadas por três hélices em cada face da folha beta ( $\alpha$ C,  $\alpha$ D e  $\alpha$ H em uma face e  $\alpha$ A,  $\alpha$ B e  $\alpha$ J na outra), formando o clássico *Rossmann-fold* (Figura 9A). <sup>139–141</sup> O domínio de ligação ao substrato é principalmente helicoidal além de duas fitas  $\beta$  paralelas curtas ( $\beta$ 7 e  $\beta$ 9). Surpreendentemente, as 4 hélices do domínio ( $\alpha$ I,  $\alpha$ K,  $\alpha$ M e  $\alpha$ N) não foram mencionadas por *MAJOR et al.*, <sup>55</sup> na sua descrição da estrutura de *At*GME embora presentes. Inclusive, globalmente os enovelamentos de *At*GME e *Md*GME são muito semelhantes com um RMSD de 0.30 Å após sobreposição dos C $\alpha$ . Além dos elementos de estrutura secundária, identificamos neste domínio três loops (A, B e C-Terminal) importantes para a ligação ao substrato em *Md*GME (**Figura 9B**).



**Figura 9-** Topologia de *Md*GME. (A) Representação em fita do monômero de *Md*GME (esquerda), mostrando um enovelamento tipo *Rossman-fold* (dominio superior) características da superfamília SDRs. O dominio de ligação ao substrato é composto principalmente de hélices-α (azul escuro na parte inferior da figura). As moléculas de NAD<sup>+</sup>, GDP-α-D-Manose e GDP-β-L-galactose ligadas são mostradas. Os átomos de carbono NAD<sup>+</sup> estão em preto, em amarelho temos os átomos de GDP-α- D-Manose (GDD) e em magenta GDP-β-L-galactose (GDC). Em NAD<sup>+</sup> os átomos de oxigênio estão demonstrados em vermelho, nitrogênio em azul e fósforo em laranja. (B) Diagrama da topologia com um enovelamento tipo *"Rossmann fold"* de *Md*GME, mostrando em verde as 7 β-fitas paralelas, em turquesa as 10 hélices α do domínio que liga a NAD<sup>+</sup>. Do lado dereito, na cor verde as 2 β-fitas paralelas e, em azul as 4 hélices α do domínio que liga ao substrato. Além disso, apresentamos em vermelho os loops A, B e C, importantes para fechar a estrutura quando o açúcar liga no sítio ativo de *Md*GME.

Fonte: Elaborada pelo autor

# 3.1.5 Sítios de ligação do cofator NAD<sup>+</sup> e substrato GDP-D-manose

A relação de GME com outras enzimas da superfamília SDR foi relatada por Major et al., 2005. <sup>55</sup> Realizaram um alinhamento estrutural com 4 homólogos de GME presentes na superfamília SDR: GMER de *E. coli* (PDB: 1BSV <sup>142</sup>); o domínio ArnA descarboxilase (PDB: 2BLL <sup>143</sup>); e UDP-galactose 4-epimerase (GalE) (PDB: 1XEL <sup>144</sup>) e RmlB (PDB: 1BXK). Usando o alinhamento, identificaram 17 resíduos conservados, dois destes resíduos foram parte da tétrade catalítica, tirosina (Y174) e lisina (K178). O terceiro e quarto aminoácido da tétrade catalítica foram encontrados como Ser/Thr (S143) e Asn. <sup>119</sup> Os outros 13 resíduos conservados estão relacionados com a ligação ao substrato ou para cumprir um papel

estrutural. Também, identificaram o resíduo N203 importante para o reconhecimento do substrato, o mesmo resíduo que está presente em *Md*GME como N202.

Para identificar os aminoácidos importantes tanto para ligação do cofator como para especificidade do substrato descrito em *At*GME, na nossa proteína *Md*GME, realizamos um alinhamento de sequências entre as duas proteínas (incluindo GME de arroz *Oryza sativa* e de uma bactéria termófila de interesse biotecnológico *Methylacidiphilum fumariolicum*) (**Figura 10**). A proteína *Md*GME apresenta os motivos  $G_{XX}G_{XX}G$  (cor vermelho) e  $Y_{XXX}K$  (cor verde), porém com uma substituição da última glicina por alanina do motivo de ligação a NAD<sup>+</sup> (**Figura 10**).

1	N-	<b>B1</b>	(αΑ ()-	
	13	34	40	55
<i>Md</i> GME	MWSTDGTDYGAFTYENLEREPYWPSI	EKLRISIT <mark>GAG</mark> G	<b>FIA</b> SHIARRLKSH	GHYIİASD
<b>MfGME</b>		-MKKALVA <mark>GAGG</mark>	FIGHHLVNFLKEP	GYWVRGVD
Atgme	GAMGTTNGTDYGAYTYKELEREQYWPSI	ENLKISIT <mark>GAG</mark> G	FIA <sup>SHIARRLKHE</sup>	GHYVIASD
0sGME	MGSSEKNGTAYGEYTYAELEREQYWPSI	EKLRISIT <mark>GAG</mark> G	FIGSHIARRLKSE	GHYIIASD
	$\overline{\mathbf{a}}\mathbf{R}$ $\mathbf{R}$	B4		( aD ()
			100	
MdGME	WKKNEHMTEDMFCHEFHLVDLRVMDNCI	LKVTKGVDHVFN	LAADMGGMGFIQS	NHSVIMYN
<b>MfGME</b>	IKEPEYEKSRSDEFYLLDLRYWGNCI	LEATKGVDEVYQ	LAADMGGIGYIS	SNHAEIAKN
AtGME	WKKNEHMTEDMFCDEFHLVDLRVMENCI	LKVTEGVDHVFN	LAADMGGMGFIQS	SNHSVIMYN
0sGME	WKKNEHMTEDMFCHEFHLVDLRVMDNCI	LKVTNGVDHVFN	LAADMGGMGFIQS	SNHSVIMYN
	$(\alpha D)$ (B5) $\alpha B$	$\alpha \mathbf{F}$	aG	aH (
	+ <sup>125</sup>	150		<b>→</b> <sup>175</sup> <b>→</b>
MdGME	TMISFNMLEAARINGVKRFFYAS	IYPEFKQLET-N	VSLKESDAWPAE	PODATGLER
<b>MfGME</b>	NILINTHMLEASYQNGVKRYFYSS <mark>S</mark> AC	LYPSYRQQSVDV	IPLKEEDAMPAD	PEEG <mark>YGWEK</mark>
AtGME	NTMISFNMIEAARINGIKRFFYAS <mark>SAC</mark>	LYPEFKQLETTN	VSLKESDAWPAE	QDA <mark>FGLEK</mark>
0sGME	NTMISFNMLEAARINGVKRFFYAS <mark>S</mark> AC	YPEFKQLET-N	VSLKESDAWPAE	QDA <mark>YGLEK</mark>
			al (	87
	200		225	
MdGME	LATEELCKHYTKDFGIECRIGRFHNIY	GPFGTWKGGREK	APAAFCRKDLTS	DKFEMW
<b>MfGME</b>	LFAEKLCQYYQEDKGIETRIARFHNVY	GPLGTYKGGREK	APAAICRKIALAH	DSSEIEVW
AtGME	LATEELCKHYNKDFGIECRIGRFHNIY	GPFGTWKGGREK	APAAFCRKAQTSI	DRFEMW
0sGME	LATEELCKHYTKDFGIECRVGRFHNIY	GPFGTWKGGREK	APAAFCRKAQTS	DRFEMW
		68	- αΚ ()-	<b>B9</b>
	250		275	
MdGME	GDGLQTRSFTFIDECVEGVLRLTKSDF	REPVNIGSDEMV	ŚMNEMAEIVLGFE	INKNLPI-H
<b>MfGME</b>	GDGKQTRSFLYIQDCVEGIYLITQSDY	PKPLNLGSEELV	TIDQLVEMTAKVA	NKNIRIRH
AtGME	GDGLQTRSFTFIDECVEGVLRLTKSDF	REPVNIGSDEMV	SMNEMAEMVLSFE	EEKKLPI-H
0sGME	GDGLQTRSFTFIDECVEGVLRLTKSDFF	REPVNIGSDEMV	SMNEMAEIILSFE	EDRELPI-H
	BO Loop B		M	
	<b>1</b> 2 300	325		350
MdGME	HIPGPEGVRGRNSDNTLIKEKLGWAPT	RLKDGLRITYF	WIKEQIEKEKAQO	MDLSIYGS
<b>MfGME</b>	NLSKPQGVRGRNSDNSKLYKITGWRPKI	FPLLEGLKLTYP	WIAERVARERNM	QGCQ
AtGME	HIPGPEGVRGRNSDNNLIKEKLGWAPN	<b>IRLKEGLRITYF</b>	WIKEQIEKEKAKO	SDVSLYGS
0sGME	HIPGPEGVRGRNSDNTLIKEKLGWAPT	<b>KLKDGLRFTYF</b>	WIKEQIEKEKTQO	SVDIAGYGS
	Loop C			
	370			
MdGME	SKVVGTQAPVQLGSLRAADGKE			
MfGME				
Atgme	SKVVGTQAPVQLGSLRAADGKE			
0sGME	SKVVSTQAPVQLGSLRAADGKE			

Figura 10 - Alinhamento de múltiplas sequências de quatro GMEs homólogos (*Md*, *Myrciaria dubia* (OK642677.1); *Mf*, *Methylacidiphilum fumariolicum* (I0K0X9); *At*, *Arabidopsis thaliana* (2C54\_A) e *Os*, *Oryza sativa* (NP\_001390897) O motivo de Glicina de ligação a NAD<sup>+</sup> (vermelho), motivo deY<sub>XXX</sub>K (verde); resíduos da tétrade catalítica cor azul, sombreado com verve e com uma estrela (vermelho), resíduo de cistina (vermelho) importante para a catálise. As α -hélices, β -fitas e loops acima indicadas, a numeração são pertencentes a *Myrciaria dubia* (*Md*GME).

Fonte: Elaborada pelo autor

# 3.1.6 Especificidade de ligação ao NAD<sup>+</sup>

O NAD<sup>+</sup> em *Md*GME está localizado no domínio *Rossmann-fold* e é estabilizado por 14 ligações de hidrogênio, uma molécula de água interagindo diretamente com o grupo fosfato do lado da nicotinamida do cofator (PN) e o fosfato do lado da adenina (PA) e 2 interações hidrofóbicas das cadeias laterais de W59 e I122. As cadeias laterais K177, Y173, D78, D58 interagem diretamente com NAD<sup>+</sup>, além de serem resíduos conservados em outras enzimas SDR estendidas de ligação a NAD<sup>+ 55</sup> (**Figura 11**). Todos os resíduos são conservados em relação a estrutura de *At*GME.



Figura 11 - Esquema de ligação de NAD<sup>+</sup> de MdGME. As interações hidrofóbicas incluem empilhamento π por W59. Distâncias e interações foram determinadas por Discovery Studio Visualizer V21.1.0 (BIOVIA, Dassault Systèmes).

Fonte: Elaborada pelo autor.

Em *Md*GME, o NAD<sup>+</sup> é encontrado na conformação "syn" com uma ligação de hidrogênio entre o nucleotídeo fosfato e a nicotinamida, a presença dessa ligação é crucial na regulação do potencial redox do par NAD<sup>+</sup> /NADH em uma enzima SDR estendida. <sup>55,145</sup> Em *Md*GME e *At*GME, o anel de adenina está ligado próximo ao motivo GxxGxxA <sup>55</sup> entre o loop que conecta a fita  $\beta$ 1 e a hélice  $\alpha$ A. A sequência do motivo GxxGxxA é incomum em SDRs estendidos por possuir uma alanina na última posição. <sup>55,146</sup> O grupo CH<sub>3</sub> da "Ala" não altera a conformação da cadeia principal, mas há a hipótese que possivelmente diminua a flexibilidade da enzima ao criar-se um núcleo hidrofóbico maior nesta região, <sup>55</sup> formado pôr os resíduos de A40, F201, I39, L98 e A140.

# 3.1.7 Especificidade de ligação do GDP-açúcar (D-manose)

Para obter a estrutura cristalográfica da *Md*GME-NAD<sup>+</sup>-GDD/GDC foi necessário incubar a proteína com seu substrato (GDP- $\alpha$ -D-manose, **GDD**). A densidade eletrônica revelou que o GDP está na posição equatorial em relação ao anel de manose na configuração D. Mas a alta resolução da estrutura (1,25 Å), mostrou também uma segunda interpretação para a densidade eletrônica (especificamente para os átomos C5' e O3' do açúcar), sugerindo ocupação parcial pelo produto GDP- $\beta$ -L-galactose (**GDC**) estrutura refinada (**Figura 12**), portanto, inclui uma mistura de substrato e produto na proporção (0.5/0.5).



Figura 12 - Diagrama de densidade de Fo-Fc não enviesada de GDP- $\alpha$ -D-manose (GDD, cor verde) e GDP- $\beta$ -L-galactose (GDC, cor cinza). A densidade eletrônica está representada em forma de rede contornada em 2,5 $\sigma$  (esquerda). As posições refinadas finais do GDD e GDC (direita). Os átomos de carbono do GDP estão em verde, os átomos de carbono da galactose em cinza e manose em verde, as moléculas de fosfato estão na cor laranja.

Fonte: Elaborada pelo autor.

Um resultado semelhante foi obtido para AtGME<sub>nativo</sub> (homólogo mais próximo), PDB=**2C59**, onde foi observado os açúcares (GDD e GDC) na estrutura e comprovando-se mediante uma análise de HPLC. <sup>55</sup> Outros complexos foram relatados também para AtGME como por exemplo o mutante K178R (PDB=**2C45**), ligado com NAD<sup>+</sup>/NADH e GDP- $\beta$ -Lgulose (**GKE**)/GDP-4-ceto- $\beta$ -L-gulose (**GKD**), obtido sem adição exógena dos ligantes. Sua densidade eletrônica é compatível com uma mistura no cristal de dois estados; GKE/NAD<sup>+</sup> e, GKD/NADH. O complexo do mutante K217A (PDB=**2C5E**) incubado com GDD, apresentou uma estrutura cristalográfica ligada com NAD<sup>+</sup> e GDD, e neste caso não foi observado o produto ou intermediário da reação, provavelmente por conta da mutação de um dois resíduo da tétrade catalítica. Observou-se um resultado semelhante com o complexo do mutante Y174F (PDB=**2C5A**) incubado com GDC, sua estrutura cristalográfica apresentou apenas NAD<sup>+</sup> e GDC ligados.

Os anéis dos açúcares (D-manose e L-galactose) de *Md*GME estão ligados próximos à região nicotinamida do NAD<sup>+</sup>. As cadeias laterais de M105, R305, C145, S143, Y173, G103 e N202 interagem diretamente com os açúcares e o O6' do açúcar interage com uma molécula de água. Por outro lado, a região GDP, ligado no domínio C terminal da proteína, faz ligações de hidrogênio com as cadeias laterais de S355, K224, Q215, P299, E300, Q240, uma molécula de água, e com os resíduos W235, F221, M234 e A220 fazem interações hidrofóbicas. As cadeias laterais dos resíduos N242 e R305 fazem uma série de interações polares com o açúcar e um dos fosfatos (**Figura 13**). Estes resíduos são conservados na estrutura cristalográfica de *At*GME (PDB: **2C59).** 



**Figura 13** - Especificidade do substrato-produto de *Md*GME. Diagrama de GDP-α-D-manose ligado a *Md*GME, as ligações de hidrogênio são mostradas como linhas tracejadas. Distâncias e interações foram determinadas por Discovery Studio Visualizer V21.1.0 (BIOVIA, Dassault Systèmes).

Fonte: Elaborada pelo autor

A sobreposição do complexo GDD/GDC de *Md*GME revelou que, apesar da inversão do anel e da dupla epimerização, todos os átomos de carbono (C1'-C'6) e oxigênio se sobrepõem, com a exceção do átomo O3'. Os átomos de O1' e O2' não fazem ligações de hidrogênio na

GDD/GDC (PDB=**2C59**). O átomo O3' interage com o mesmo resíduo de G103 em ambos os açúcares, com uma pequena diferença por causa da conformação dos compostos. O átomo O4 forma uma ligação de hidrogênio com a S143 em ambos os casos enquanto C145 e Y173 participam de ligações de hidrogênio fracas. O átomo O5' interage com uma molécula de água e com R305.

Estudos da atividade enzimática dos mutantes K217A, K178R, Y174F e C145A de *At*GME, revelaram sua importância na atividade catalítica da enzima <sup>55</sup>. Os resíduos K178R e Y174F foram os mais importantes, eles formam parte da tétrade catalítica de *Md*GME (**Figura 14**). A superposição das estruturas cristalográficas de *At*GME<sub>nativo</sub> e *Md*GME mostra o sítio catalítico conservado, além dos resíduos R242 e N202, conservados na estruturas de GMER, RmIB, GalE e no domínio ArnA descarboxilase <sup>55</sup> (**Figura 14**).



Figura 14 – Sítio ativo de *Md*GME. Mostra a sobreposição da estrutura de *Md*GME-NAD<sup>+</sup>-GDD/GDC com *At*GME (PDB: 2C59). NAD<sup>+</sup> (cor verde), GDD (cor verde) e GDC (cor cinza). A tétrade catalítica, Ser143, Tyr173, Lys177 e Asn119 para *Md*GME (cor cinza) e para *At*GME (cor turquesa), o resíduo de C145, importante para a catálise (cinza *Md*GME e azul claro *At*GME). Resíduos R242, N202 são conservados em outras SDRs.

Fonte: Elaborada pelo autor

# 3.1.8 Mecanismo geral de GME

O mecanismo da superfamília SDR e, portanto, também das GMEs, é a transferência de um hidreto entre o substrato e o cofator  $(NAD(P)^+)$  ligado à enzima. <sup>41</sup> Em contexto, *Md*GME começaria a reação com uma oxidação do C4' de GDP-Manose em parceria com o cofator NAD<sup>+</sup> e, o resíduo de Tyr173 do sítio catalítico (Y<sub>XXX</sub>K) que atuaria como um **ácido catalítico** para auxiliar a desprotonação. <sup>41</sup> A oxidação inicial de GME em geral resulta com o C4' em

um intermediário ceto transitório, que funciona apenas para diminuir o pKa dos prótons nos carbonos vizinhos (C3' e C5'). Subsequentemente, um ácido catalítico adicional (**Cys145**) e uma base (**Lys177**) realizam a desprotonação e a reprotonação em C5' e C3'. O **enolato** que é criado durante este processo é estabilizado por ligações de hidrogênio com a Tyr173, bem como com uma Ser adicional (S143) da tétrade catalítica. <sup>41</sup>

As observações estruturais de *At*GME e o novo produto (GDP-D-altrose) encontrado recentemente na reação de GME, fizeram com que a reação de epimerização da enzima tenha sido postulada com uma dupla rota de reação: C5'-antes-de-C3' e C3'-antes-de-C5'. <sup>41,55,60</sup> Uma visão geral esquemática do mecanismo é apresentada na **Figura 15**. Os diferentes ceto-intermediarios (GDP-4-ceto- $\beta$ -L-Galactose, GDP-4- ceto- $\beta$ -L- Gulose e GDP-4- ceto- $\alpha$ -D-Altrose) poderão ser reduzidos em GDP- $\beta$ -L-Galactose (C3',5'-epímero do GDP-Manose) e dois subprodutos: GDP- $\beta$ -L-Gulose (epímero C5') e GDP-D-Altrose (epímero C3). Na primeira rota C5'-antes-de-C3' (cor verde, **Figura 15**), previram que uma inversão do anel ocorre durante a primeira epimerização (C5') e que um intermediário é provável para a segunda epimerização (C3'), sugerindo um movimento C4' antes e depois da segunda epimerização. <sup>41,55</sup>



**Figura 15** - Mecanismo da GDP-manose 3',5'- epimerase. A reação de epimerização resulta em equilíbrio entre GDP-α-D-Manose, GDP-β-L-Galactose, GDP-β-L-Gulose e GDP-α-D-Altrose

Fonte: Adaptada de BEERENS et al. 41

A rota C3'-antes-de-C5' (cor vermelho, **Figura 15**), inclui também um movimento C4', que é necessário para o posicionamento do substrato e para permitir a abstração do segundo próton em C5'. <sup>41</sup> A confirmação de que ambas as rotas na enzima GME são possíveis, a semelhança estrutural e mecanicistas com a enzima GMER (GDP-4-ceto-6-desoxi-D-manose-

3',5'-epimerase/4-redutase) sugerem uma rota de epimerização dupla para GMER, pelo fato, de ambas as enzimas empregarem resíduos semelhantes para alcançar a dupla epimerização C3',5',ou seja, Cys/Lys (Arg) e Cys/His (Lys), <sup>41,147</sup> mas, mesmo tendo os resíduos semelhantes e pertencer à mesma superfamília (SDRs), o mecanismo de catálise é bem diferente. <sup>55</sup> Neste contexto, GME é uma enzima relevante que catalisa três reações químicas distintas em um sítio ativo (oxidação, epimerização e redução) e, portanto, o nome mais correto pelas funções que faz, seria GDP-manose 4-oxidase/3,5-epimerase/4-redutase. <sup>55</sup>

#### 3.1.9 Estrutura de *Md*GME-NAD<sup>+</sup>

Os SDRs em geral possuem um grupo de resíduos conservados responsáveis pela ligação de NAD(P). <sup>137,141</sup> Neste trabalho a segunda estrutura cristalográfica resolvida foi *Md*GME em complexo com NAD<sup>+</sup> (*Md*GME-NAD<sup>+</sup>). Esta é a primeira estrutura de GME a ser resolvida com apenas um ligante, uma vez que as 4 estruturas cristalográficas de *At*GME foram resolvidas em complexo com NAD<sup>+</sup> e diferentes açúcares (GDD, GDC, GKE e GKD). A estrutura de *Md*GME-NAD<sup>+</sup> foi resolvida por substituição molecular usando a estrutura de *Md*GME-NAD<sup>+</sup>-GDD/GDC. A estrutura apresentou 4 moléculas na unidade assimétrica, porém sem apresentar os contatos que estabilizam o dímero fisiológico. Após análises com o programa PISA, <sup>133-134</sup> e aplicando operações de simetria, identificamos o dímero que apresentava os contatos fisiológicos (**Figura 16**).



Figura 16 - Estado fisiológico da enzima MdGME-NAD<sup>+</sup>. (A) Conteúdo da unidade assimétrica da enzima MdGME-NAD<sup>+</sup>, (B) Aplicação de simetria para achar o estado fisiológico da enzima, (C) Zoom de um dímero fisiológico da enzima MdGME-NAD<sup>+</sup>.

Fonte: Elaborada pelo autor

# **3.1.10** Estado aberto e fechado de *Md*GME

Uma sobreposição das estruturas cristalográficas para os complexos *Md*GME-NAD<sup>+</sup> e *Md*GME-NAD<sup>+</sup>-GDD/GDC, com um RMSD=0,901 Å é mostrada na (**Figura 17**). Como pode ser visto, uma característica importante é o fato de que a enzima *Md*GME sofre uma mudança conformacional em um dos domínios funcionais (domínio de ligação ao substrato) após a ligação do GDP- $\alpha$ -D-manose. A sobreposição de *At*GME (PDB:**2C5A**) com *Md*GME-NAD<sup>+</sup>, produz um RMSD =1,050, a divergência é semelhante dos resultados obtidos para *Md*GME-NAD<sup>+</sup>-GDD/GDC. Esta é uma mudança substancial na estrutura como resultado da ligação do substrato, as divergências observadas se definem no câmbio conformacional que ocorre no domínio de ligação ao GDP-açúcar. As principais mudanças são observadas nas hélices  $\alpha$ M,  $\alpha$ K,  $\alpha$ I e as fitas  $\beta$ 7 e  $\beta$ 9 que sofrem um desvio quadrático médio notavelmente alto de 3.8 e 12.9 Å, respectivamente. Além disso, o aparecimento da hélice  $\alpha$ N e o ordenamento dos loops A, B e C (chamados assim, para uma melhor compreensão) em MdGME-NAD<sup>+</sup>-GDD/GDC são diferenças significativas. Estas regiões apresentam resíduos importantes que participam diretamente na ligação dos GDP-açúcares. Estes resultados sugerem que o NAD<sup>+</sup> é ligado fortemente à enzima, mas na ausência do substrato o domínio da ligação ao substrato é levemente desordenado, como observado para a hélice  $\alpha$ N e loops A, B e C. Estas regiões ficam reordenados quando o GDP-açúcar se liga e, portanto, o mecanismo de abertura e fechamento de GME seria dependente de GDP-açúcar. Nossos dados, e particularmente a determinação da primeira estrutura sem um substrato/produto ligado, permitiu visualizar, os detalhes desta variação que parecem ser significativos para o mecanismo conformacional da enzima, essencial para realizar uma reação catalítica competente.



Figura 17- Os estados de abertura e fechamento de MdGME. (A e C) representam a forma apo de MdGME-NAD na cor rosa mostrando o estado aberto, (B e D) representam a forma holo de MdGME -NAD-GDD-GDC (colorida conforme a Fig. 9) mostrando o estado fechado. O círculo e a flecha mostram a sítio de abertura e fechamento, a mudança é caracterizada pelo contato entre o Loop-C (Vermelho, não observado na estrutura apo) o domínio de ligação ao açúcar, apresentam um cambio conformacional nas hélices e fitas. A superposição entre as formas MdGME apo e holo mostra um "close-up" das regiões do domínio de ligação ao açúcar.

Fonte: Elaborada pelo autor

Os resultados observados no mecanismo de abertura e fechamento da enzima MdGME, sugerem que a ligação ao NAD<sup>+</sup> pode ter um papel importante para estabilizar a interface de dimerização, além, de sua importância para a funcionalidade competente da enzima. Isso porque é o domínio de ligação ao NAD que faz a interface de contato entre subunidades e não se altera com a mudança conformacional descrita acima. Nesse contexto, identificamos os resíduos mais importantes para a dimerização da enzima MdGME (**Tabela 3**).

Aminoácidos	Estrutura secundaria	Aminoácidos	Estrutura secundaria	Aminoácidos	Estrutura secundaria
MET 82	αC	ARG 131	αD	HIS 186	αH
HIS 112	αD	TRP 165	Loop aG- aH	TYR 187	αH
SER 113	αD	PRO 166	Loop aG- aH	ASP 190	αH
MET 116	αD	ALA 167	Loop aG- aH	PHE 191	αH
TYR 117	αD	GLU 168	Loop aG- aH	THR 360	Loop C
THR 120	αD	ASP 171	Loop aG- aH	GLN 361	Loop C
MET 121	αD	GLU 176	αH	ALA 362	Loop C
PHE 124	αD	ALA 179	αH	PRO 363	Loop C
ASN 125	αD	GLU 182	αH	LEU 366	Loop C
GLU 128	αD	LEU 183	αH		

**Tabela 3** – Resíduos envolvidos na dimerização de GME. As interações foram determinadas por Ligplot. <sup>148</sup>

Fonte: Elaborada pelo autor.

Na **Figura 18**, pode ser observar que são as hélices  $\alpha D \in \alpha H$  de uma face do *Rossmann-fold* que formam a maior parte da interface de contacto entre as duas subunidades do dímero junto a hélice  $\alpha C$  e o loop que conecta a hélice  $\alpha G$ - $\alpha H$  e o loop C. <sup>41,55</sup> Para visualizar as regiões da interface, o dímero foi aberto como um livro e o monômero foi separado para mostrar os resíduos que tem contacto na interface de dimerização (**Figura 18-lado direito**).

Os padrões gerados ao mapear os resíduos de contato de MdGME-NAD<sup>+</sup>-GDD/GDC, são semelhantes aos observados para AtGME (PDB=**2C5A**) e MdGME-NAD<sup>+</sup>. A diferença mais notória é a perda do contato do loop C em MdGME-NAD<sup>+</sup>. Este resultado nos sugere que o papel mais importante do loop C é o empacotamento da estrutura e sua estabilidade, quando não tem ligado o GDP-açúcar observe-se sua flexibilidade sugerindo que a enzima é menos estável. Antes de nosso trabalho, todas as estruturas resolvidas de GME mostraram sua conformação mais ordenada por conta da presença do substrato/produto junto com o cofator ligado. A grande conservação estrutural entre *At*GME e *Md*GME sugere que o mecanismo de abertura e fechamento em torno do substrato talvez seja comum entre eles e talvez se estenda para os demais integrantes da família.



**Figura 18** - As interfaces da dimerização de *Md*GME ligado com o GDP-açúcar e NAD<sup>+</sup>. O monômero foi separado e girado em 90° o em torno do eixo vertical, para expor os resíduos da interface (branco).

Fonte: Elaborada pelo Autor

# 3.2 Subcapítulo: L-galactose desidrogenase (GDH)

#### 3.2.1 O estado oligomérico da L-galactose desidrogenase

Conforme descrito na metodologia, a expressão heteróloga e purificação de L-galactose desidrogenase de camu-camu MdGDH e espinafre SoGDH, foi realizada através de duas etapas. A primeira, constituída por uma cromatografia de afinidade ao níquel. As bandas no gel de SDS-PAGE (**Figura 19A**) mostram a proteína solúvel entre as massas moleculares de 31kDa e 45kDa, sendo que a massa molecular teórica de MdGDH é de 39,30 kDa e, para SoGDH de 37,55 kDa. Este resultado mostra que MdGDH purifica com uma baixa quantidade de contaminante nas frações de eluição E1 e E2 enquanto, SoGDH se apresenta praticamente pura nas frações de eluição E2, E3 e E4. Além disso, SoGDH apresentou um rendimento maior que da proteína MdGDH, isso poderia estar relacionado com o vetor [pET28a(+)] que foi usado para a expressão heterologa da proteína recombinante.



Figura 19 – Expressão e purificação de GDHs. A) Cromatografia de Afinidade de MdGDH (cima) e SoGDH (baixo), M:Marcador de massa molecular, AI: antes de induzir com IPTG, DI: depois de induzir com IPTG, P: pellet-fração Insolúvel, S: fração solúvel, NL: proteínas que não interagem com o níquel imobilizado na resina, L1: lavagem com tampão, E1-E5: eluições com 25, 50, 100, 250 e 500 mM de imidazol. Eletroforese em gel SDS-PAGE 12%. B) Cromatografia de Exclusão de Massa Molecular (SEC) de GDHs, com picos homogêneos.

Fonte: Elaborada pelo autor

Na segunda etapa realizou-se uma cromatografia de exclusão de massa molecular, onde as eluições que apresentaram uma banda única na etapa 1, foram concentradas e injetadas na coluna HiLoad Superdex 200 16/60. O cromatograma de purificação (**Figura 19B**), apresenta picos bem definidos para ambas as proteínas e, suas purezas foram analisadas em gel de SDS-PAGE 12%, mostrando uma única banda.

Para dar continuidade às próximas etapas de caracterização, foi importante observar a homogeneidade e estado oligomérico das enzimas recombinantes das proteínas *MdGDH* e *SoGDH*. Para isso, usamos a cromatografia de exclusão de massa por tamanho acoplada aos perfis de espalhamento de luz multi-ângulo (SEC-MALS) como técnica absoluta para determinar a massa molecular de nossas proteínas em solução, assim como a homogeneidade dela. Na **Figura 20** apresentam-se ambas as enzimas recombinantes eluidas aproximadamente no mesmo volume, com picos predominantemente homogêneos e, com a massa molecular teórica esperada para um monômero.



Figura 20 - Cromatograma de SEC-MALS de GDHs. Tem-se em vermelho curvas correspondentes à mudança no índice de refração diferencial normalizado, em azul a intensidade de luz espalhada normalizada a 90° para MdGDH e em turquesa para SoGDH. A linha tracejada horizontal preta representa a massa molecular determinada. O MW determinada de MdGDH e SoGDH (36,4 kDa e 35,0 kDa, respetivamente) o que corresponde ao MW teórica esperado para um monômero de GDH.

Fonte: Adaptada de VARGAS et al. 132

Até o momento estudos relacionados ao estado oligomérico em GDH exibiu variação, dependendo do organismo e da fonte de isolamento. Por exemplo, a partir da purificação do eixo embrionário de *Pisum sativum* (ervilha, *Ps*GDH) foi relatada como uma proteína tetramérica. <sup>61</sup> Por outro lado, purificados das folhas de *Spinacia oleracea* (spinafre, *SoGDH*) e *Actinidia deliciosa* (kiwi, *AdGDH*), foram descritos como homodímeros e monômeros, respectivamente. <sup>62-63</sup> Enquanto estudos com enzimas recombinantes, *Arabidosis thaliana* (*At*GDH) e *Oryza sativa* (arroz, *Os*GDH) foram relatadas como homodiméricas e monoméricas, respectivamente. <sup>61,64</sup> Embora existam diferenças nas técnicas usadas para purificação de proteínas de fontes naturais ou heterólogas, acredite-se que isso não deveria afetar o estado oligomérico da enzima. Além disso, os estudos que relataram a enzima como dímeros ou tetrâmeros forneceram dados limitados para fundamentar essas alegações <sup>61-62</sup> e os estudos que relataram o estado monomérico para GDH usaram apenas análise cromatográfica de exclusão molecular de tamanho. <sup>63-64</sup>

O estado monomérico dos GDHs estudados aqui com SEC-MALS é consistente com a inclusão de GDH como parte da superfamília de proteínas aldo-ceto redutases (AKR), <sup>61</sup> composto por 16 famílias distribuídas em todos os filos. Essas enzimas catalisam reações de oxidação-redução em substratos carbonílicos, como por exemplo: aldeídos de açúcar, ceto-

esteróides, quinonas, etc. Geralmente são monômeros cujo enovelamento é dominada por um motivo TIM-barril ( $\beta/\alpha$ )8, são caracterizadas por uma região de ligação de cofator conservada NAD(P)H, possuem estruturas de loop variáveis que definem a especificidade pelo substrato e tem uma tétrade catalítica.

Das 16 famílias de AKRs, apenas três foram relatas como diméricas: Aflatoxina dialdeído redutase de fígado de rato (AKR7A1), xilose redutase dependente de NAD(P)H de *Candida tenuis* (AKR2B5) e a proteína supressora de auxotrofia de tirosina de *Escherichia coli*. As estruturas cristalográficas dessas proteínas mostram uma interface de contato conservada para a formação de dímeros, que está ausente das enzimas aqui descritas. Isso será discutido posteriormente a partir estrutura cristalografica de *SoGDH*, o qual sugere que GDH é, em geral, uma enzima monomérica.

#### 3.2.2 Enovelamento e estabilidade térmica de MdGDH e SoGDH

Avaliamos o estado de enovelamento das proteínas recombinantes mediante o monitoramento do espectro de CD (a cada 2°C) em função do aumento da temperatura atraves de um gradiente de 10 a 80°C. A **Figura 21A e C** mostram os espectros de CD das proteínas *So*GDH e *Md*GDH respectivamente em temperaturas onde houveram mudanças do sinal. Na temperatura de 20°C, podemos observar mínimos negativos em 222 e 208 nm, que representam espectros característicos de proteínas que contêm elementos de estrutura secundária  $\alpha/\beta$ . Os espectros a uma temperatura de 50°C mostram uma diferença significativa, que está relacionada com a perda de elementos de estrutura secundária tanto em *So*GDH, como em *Md*GDH. A curva de desnaturação térmica, onde o sinal de espectro CD foi monitorado a 222nm, mostra a temperatura de melting (T<sub>m</sub>) obtida de 52,0 ± 0,2 para *So*GDH e 55,6 ± 0,1 para *Md*GDH após um ajuste de *Boltzman* no sinal detectado (**Figura 21B e D**). Estes resultados mostram uma boa estabilidade térmica para ambas as enzimas recombinantes e um comportamento semelhante.



Figura 21 - Espectro de CD de GDHs. (A e C) Mostram os espectros de CD obtidos em diferentes temperaturas de SoGDH e MdGDH respectivamente. (B e D) Curva de desnaturação térmica para SoGDH e MdGDH respectivamente, Tm entre 52,0°C e 55,6°C (caixa vermelha).

Fonte: Elaborada pelo autor

# 3.2.3 Parâmetros Cinéticos de GDHs

A enzima recombinante de camu-camu *Md*GDH apresenta cinética tipo Michaelis-Menten usando como seu substrato L-galactose (L-gal) e excesso de NAD<sup>+</sup> (0,2 mM), com valor de  $K_m$  de 0,21 mM, constante catalítica ( $k_{cat}$ ) de 4,26 s<sup>-1</sup> e uma eficiência catalítica ( $k_{cat}/K_m$ ) de 20,67 mM<sup>-1</sup>s<sup>-1</sup> (**Figura 22A**), com valor de pH ótimo de 7,0. Por outro lado, a enzima recombinante do espinafre (*So*GDH) apresentou um K<sub>m</sub> de 0,13 mM (**Figura 22C**), próximo ao relatado para *So*GDH purificada a partir de folhas de espinafre (0,12 mM). <sup>62</sup> Consequentemente, de acordo com relatos recentes, <sup>61–63</sup> o valor de K<sub>m</sub> para GDH de diferentes espécies de plantas varia de 0,08 a 0,43 mM.



Figura 22 - Propriedades cinéticas de camu-camu MdGDH e espinafre SoGDH. Cinética do tipo Michaelis-Menten foi observada para MdGDH (A) e SoGDH (C) com valores de K<sub>m</sub> de 0,206 e 0,128 mM, respectivamente (valores médios de cinco medições independentes com barras de erro correspondentes). As curvas de atividade em função de pH são mostradas para MdGDH (B) e SoGDH (D).

Fonte: Adaptada de VARGAS et al. 132

Por outro lado, observamos um pH ótimo de 7,0 para *So*GDH, semelhante ao *Md*GDH (**Figura 22B e D**), enquanto *MIEDA et al.* <sup>62</sup> relatou um valor de 9,25. Para *So*GDH observamos um  $k_{cat}$  de 1,2 s<sup>-1</sup> e uma eficiência catalítica de 9,1 mM<sup>-1</sup>s<sup>-1</sup>. Em geral, as duas enzimas mostram parâmetros cinéticos que são semelhantes e bem dentro de uma ordem de magnitude dos valores relatados para outros espécies (**Tabela 4**). No entanto, a discrepância nos valores relatados para o pH ótimo de *So*GDH é difícil de explicar neste momento, mas é importante notar que o artigo de *MIEDA et al.* <sup>62</sup> não fornece detalhes experimentais sobre como o valor de pH ótimo de 9,25 foi determinado. Além disso, os autores citam apenas o uso do tampão Tris, que apresenta uma faixa tamponante entre 7 e 8, não resistindo às variações de pH para cima ou para baixo desses valores. Mais estranho ainda, apesar de ter citado o pH 9,25 como o pH ideal para GDH de espinafre, os estudos cinéticos foram realizados em tampão Tris 100 mM em pH 7,5.

Tabela 4 - Parâmetros cinéticos para GDHs.

Enzimas	K <sub>m</sub> (mM)	k <sub>cat</sub> (s <sup>-1</sup> )	$\frac{k_{cat}/K_m}{(mM^{-1}s^{-1})}$	K <sub>i</sub> (mM)	pH ótimo
Camu-Camu MdGDH <sup>a</sup>	0.206	4.261	20.67	1.2	7.0
Espinafre SoGDH <sup>a</sup>	0.128	1.158	9.06	0.5	7.0
Arabidopsis AtGDH <sup>b</sup>	0.085	N/D	N/D	N/D	8.5-9
Pea <i>Ps</i> GDH <sup>b</sup>	0.430	N/D	N/D	N/D	8.5-9
Kiwi AdGDH <sup>c</sup>	0.300	N/D	N/D	0.5	9.0
Espinafre SoGDH <sup>d</sup>	0.116	N/D	N/D	0.1	9.25

<sup>a</sup>GDH parâmetros de acordo com este estudo.

<sup>b</sup>GDH parâmetros de acordo com GATZEK et al. <sup>61</sup>

<sup>c</sup>GDH parâmetros de acordo com LAING *et al.* <sup>63</sup>

<sup>d</sup>GDH parâmetros de acordo com MIEDA et al.<sup>62</sup>

Fonte: Adaptada de VARGAS et al. 132

No estudo de *MIEDA et al.*<sup>62</sup> também sugeriram a regulação por retroalimentação da GDH pelo AsA, o produto final da via biossintética D-manose/L-galactose, por inibição competitiva. Conforme este relato, o AsA a 1 mM inibe até 41% da atividade de *So*GDH com um K<sub>i</sub> de 0,13 mM. De forma semelhante, *At*GDH também foi inibida por AsA.<sup>61</sup> No entanto, uma análise mais detalhada do efeito inibitório do AsA sobre a enzima do kiwi (*Ad*GDH) mostrou que a inativação pelo AsA é lenta e irreversível, possivelmente atribuída ao dano oxidativo de um resíduo de aminoácido chave no sítio ativo <sup>63</sup>, mas o resíduo especifico não foi identificado.

O camu-camu é uma planta que armazena altas concentrações de AsA (2g/100g de polpa) em seus frutos. <sup>63</sup> No contexto deste fruto acumular uma alta concentração de AsA, o *Md*GDH poderia ser potencialmente resistente à regulação por feedback do AsA. Para testar essa hipótese, realizamos ensaios de inibição de *Md*GDH e *So*GDH recombinantes com concentrações de AsA de até 4,5 mM (**Figura 23**). AsA a 1,5 mM, 3 mM e 4,5 mM inibiu a atividade de *Md*GDH em 53%, 71% e 79%, respectivamente (K<sub>i</sub> = 1,2 mM). Nas mesmas concentrações de AsA, a atividade de *So*GDH foi inibida em 82%, 89% e 92%, respectivamente (K<sub>i</sub> = 0,5 mM) (**Figura 23A e B**). Estes resultados indicam que camu-camu *Md*GDH tem uma tolerância ligeiramente maior à inibição por AsA do que espinafre *So*GDH e kiwi *Ad*GDH.



Figura 23 - A inibição de MdGDH e SoGDH por AsA em tampão Tris-HCl 100 mM (A e B) ou Tris-HCl 300 mM (C e D) é mostrada. O efeito inibitório do AsA cai drasticamente utilizando-se um tampão com concentração mais alta.

Fonte: Adaptada de VARGAS et al. 132

Embora o AsA seja apenas um ácido fraco, as concentrações utilizadas nos ensaios de inibição podem levar a alterações significativas no pH do tampão de reação (Tris-HCl 100 mM, pH 7,0). Considerando este efeito e sabendo que mudanças no pH alteram a atividade da GDH (**Figura 22B e D**), realizamos medições de pH do tampão de reação com cada uma das concentrações de AsA utilizadas no ensaio. Observamos que ao aumentar a concentração de AsA, o pH da solução diminuiu. Por exemplo, usando uma solução de reação com Tris-HCl pH 7,0 a 100 mM, o pH caiu para 5,9 na presença de 4,5 mM AsA. As curvas clássicas de dependência de pH mostradas nas **Figura 23C e D** indicam que esta queda de pH de > 1 unidade de pH teria um impacto significativo na atividade enzimática, da ordem que observamos experimentalmente. Isso sugere que a inibição observada tanto no camu-camu quanto no espinafre GDH pode ser efeito da mudança do pH e não de uma inibição por
retroalimentação própria pelo AsA. Para validar nossas hipóteses, ensaios de inibição foram realizados na presença de um tampão mais forte (300 mM Tris-HCl) onde o pH do tampão de reação não se seria drasticamente afetado pelas concentrações de AsA (**Tabela 5**). Nestas condições, o AsA não apresentou efeito inibitório sobre a atividade de *So*GDH (**Figura 23D**). No entanto, no caso de *Md*GDH, a inibição leve persistiu com um K<sub>i</sub> estimado de 2,1 mM, esse efeito pode estar relacionado ao próprio tampão de reação, pois ao comparar os valores de K<sub>m</sub> da enzima na ausência de AsA nos tampões de 100 e 300 mM, varia de 0,206 mM a 0,650 mM, respectivamente. No entanto, o efeito produzido pelo AsA sobre a atividade enzimática parece ser em grande parte devido à queda do pH no caso de ambas as enzimas. Vale ressaltar que *MIEDA et al.* <sup>62</sup> também realizaram seu estudo usando tampão Tris na concentração de apenas 100 mM. Os dados obtidos neste trabalho questionam as conclusões do estudo de *MIEDA et al.* <sup>62</sup> que sugerem a inibição por retroalimentação negativa pelo produto da via (AsA). Não encontramos nenhuma evidência para isso no presente estudo, refutando o GDH como um potencial mecanismo regulador da via D-manose/L-galactose.

Tabela 5 –	Medidas	de pH no	o tampão	de reação
------------	---------	----------	----------	-----------

AsA (mM)	100 mM Tris-HCl	300 mM Tris-HCl		
0	7.02	7.11		
0.75	6.87	7.10		
1.5	6.78	7.09		
3	6.60	7.03		
4.5	5.90	6.99		

Fonte: Elaborada pelo autor.

# 3.2.4 Descrição geral da estrutura de SoGDH

Neste trabalho abordamos diversas tentativas de cristalização (com cofator, sem cofator e com AsA como possível inibidor) para camu-camu *Md*GDH e espinafre *So*GDH, das quais as condições mais favoráveis que resultaram em cristais com qualidade de difração de raios-X estão ilustradas na **Figura 24**. Dados de difração de qualidade só foram coletados para *So*GDH (sem cofator) e *So*GDH ligada a seu cofator NAD<sup>+</sup> (*So*GDH-NAD<sup>+</sup>). As estruturas foram refinadas a resoluções de 1,40 e 1,75 Å, respectivamente. A boa qualidade dos modelos se reflete nos valores de R<sub>work</sub> e R<sub>free</sub>, para *So*GDH com valores de R<sub>work</sub>= 19,42% e R<sub>free</sub> = 20,95%, com 97,76% de aminoácidos nas regiões mais favorecidas do espaço Ramachandran.

No caso de SoGDH-NAD<sup>+</sup> os valores equivalentes são = 18,37%, 21,48% e 98,08% respectivamente. Um resumo dos parâmetros da coleta e estatística de refinamento são mostrados na **Tabela 6.** 

	SoGDH	SoGDH-NAD <sup>+</sup>		
Fonte de X-ray	Sirius MANACA	Diamond I03		
Detector	Pilatus 2M	Eiger2 XE		
Parâmetros da cela: a, b, c (Å)	51.72, 55.01, 61.79,	49.99, 54.97, 63.22		
Parâmetros da cela: $\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 112.10, 90.00	86.84, 69.13, 84.14		
Grupo especial	$P2_1$	<i>P</i> 1		
Resolução (Å)	39.69 - 1.40 (1.42 -	39.73 - 1.75 (1.78 -		
Kesolução (A)	1.40)	1.75)		
λ (Å)	1.3236	0.97622		
Multiplicidade	12.7 (12.1)	3.4 (3.3)		
$R_{\rm pim}$ (all I+ & I-) (%)	4.0 (48.5)	6.9 (38.1)		
CC(1/2)	0.998 (0.643)	0.958 (0.706)		
Completeza (%)	99.6 (97.6)	87.5 (40.8)		
Reflexões	804980 (36548)	190355 (4645)		
Reflexões únicas	63165 (3024)	55241 (1416)		
$< I/\sigma(I) >$	14.0 (2.7)	7.5 (1.8)		
<i>R</i> (%) *	19.42	18.37		
$R_{\text{free}}$ (%) *	20.95	21.48		
Número de átomos: proteína	2421	4845		
Número de átomos: Água	250	533		
Número de átomos: ligante	0	88		
B (Å <sup>2</sup> )	20.08	14.07		
Erro de coordenada (basado no- ML) (Å)	0.14	0.19		
Erro de fase (°)	21.96	24.50		
Ramachandran favorecida (%)	97.76	98.08		
Ramachandran permitida (%)	2.24	1.92		
Pontuação de confrontos de todos os átomos	1.45	2.87		
Comprimento de ligações (RMSD) (Å)*	0.009	0.006		
Ângulo de ligações (RMSD) (°)*	1.043	0.790		
Estruturas depositadas no PDB	7SMI	7SVQ		

Tabela 6 - Coleta de dados de raios-X e estatísticas de refinamento de estrutura de GDHs.

Fonte: Adaptada de VARGAS et al. 132

Para a proteína *Md*GDH as tentativas de cristalização não foram bem-sucedidas e os cristais obtidos se resumem em apenas uma única condição (SG1-G6) com cristais muito pequenos (**Figura 24D**). No entanto, as tentativas de melhorar o tamanho dos cristais foram bem-sucedidas (**Figura 24E e F**), mas os mesmos difrataram a baixa resolução (3-4 Å). O melhor conjunto de dados que obtido foi de 2,7 Å, mas os dados foram anisotrópicos, o qual ainda não permitiram determinar sua estrutura cristalografia.



Figura 24 - Cristais das proteínas GDHs. A, B e C são cristais de *So*GDH e D, E e F são cristais de *Md*GDH.
(A) Condição A4 (2 M Sulfato de Amônio, 0,1 M Bis-Tris pH 6.5) do kit de cristalização Index<sup>TM</sup>,
(B) Condição A3 (2 M Sulfato de Amônio, 0,1 M Bis-Tris pH 5.5) do kit de cristalização Index<sup>TM</sup>,
(C) Condição G6 (2 M Sulfato de Amônio, 0,1 M Bis-Tris pH 5.5, 25% w/v de polietilenoglicol 3,350) do kit de cristalização Index<sup>TM</sup>.
(D) condição G6 (0,5 M de sulfato de amônio, 1,0 M sulfato de Lítio, 0,1 M Citrato de sódio pH 5.6 do kit de cristalização SG1. (Condição E e F) Modificações da condição de G6 no pH (0,5 M de sulfato de amônio, 1,0 M sulfato de Lítio, 0,1 M Citrato de soluções foram realizadas com reagentes de grado molecular de 99% puro e se-utilizou a concentração de10 mg/mL de proteína (com e sem cauda de Histidina).

Fonte: Elaborada pelo autor

Por outro lado, uma estudo preliminar de raios-X de *Os*GDH de arroz (*Oryza sativa*) foi relatado em 2013, mas a estrutura permaneceu sem resolução devido a dificuldades com a substituição molecular. <sup>64</sup> No entanto, em 2021 os mesmos autores depositaram duas estruturas para a enzima, uma na forma apo a 1,2 Å (*Os*GDH [PDB: 7EZI]) e outra na forma de um complexo com o cofator NAD<sup>+</sup> a 1,8 Å (*Os*GDH-NAD<sup>+</sup> [PDB: 7EZL]). Porém, nenhuma publicação descrevendo os detalhes dessas primeiras estruturas apareceram na literatura. No momento em que as estruturas relatadas aqui foram resolvidas, as estruturas

homólogas do arroz ainda não estavam disponíveis e inicialmente tivemos dificuldades semelhantes em encontrar uma solução de substituição molecular. No entanto, a estrutura da enzima apo (*So*GDH) foi finalmente resolvida usando MoRDa<sup>114</sup>, um pipeline para substituição molecular automatizada com base em um banco de dados de domínio de proteína derivado do PDB.<sup>114</sup> A estrutura do complexo ligado ao cofator (*So*GDH-NAD<sup>+</sup>) foi resolvida por substituição molecular usando a estrutura apo como modelo de busca com o programa *Phaser* no pacote *Phenix*.<sup>118</sup> A estrutura cristalográfica da proteína *So*GDH possui uma molécula na unidade assimétrica com um enovelamento ( $\beta/\alpha$ ) 8-barril, como observado na superfamília AKR (**Figura 25A**).<sup>149</sup> Possui 8 fitas  $\beta$  paralelas que são alternadas com 8  $\alpha$ -hélices que correm antiparalelas em relação às fitas, formando o enovelamento clássico do tipo barril (**Figura 25**). No terminal N-terminal, há um loop  $\beta$  adicional ( $\beta$ 1 e  $\beta$ 2) que forma o fundo do barril.



Figura 25 - Topologia de SoGDH. A) Representação em 3D. B) Diagrama da topologia para a estrutura do SoGDH que mostra um enovelamento (β/α) 8-barril (TIM-barril), característica da superfamília AKR. Este enovelamento tem 8 β-fitas (amarelas) intercaladas por 8 α-hélices (turquesa). Na parte inferior do barril, no terminal N da cadeia polipeptídica, há um loop β composto pelas fitas β1 e β2. Além disso, SoGDH apresenta duas hélices (H1 e H2) externas ao barril, que juntamente com os Loop-A, Loop-B e Loop-C são importantes para a classificação dos membros de AKRs.

Fonte : Adaptada de VARGAS et al. 132

O  $(\beta/\alpha)$ 8-barril é conservado na superfamília AKR, porém os loops que ligam as fitas e hélices, bem como a presença de hélices externas ao barril, podem variar bastante e são importantes para a classificação dos AKRs na classe correta entre as 16 famílias diferentes. <sup>150</sup> Os loops A ( $\beta$ 6- $\alpha$ 4), B ( $\beta$ 9- $\alpha$ 7) e C (a região C-terminal) são os mais importantes porque estão relacionadas à ligação do cofator e/ou especificidade do substrato.<sup>149</sup>

A **Figura 26** mostra um alinhamento de sequências incluindo cinco representantes de GDHs de diferentes espécies de plantas juntamente com sete AKRs de diferentes famílias que possuem estruturas 3D conhecidas e disponíveis. As GDHs possuem um Loop-A curto, semelhante ao observado nas famílias AKR3, AKR5, AKR7 e AKR11, mas diferente quando comparado a AKR1, AKR2 e AKR4. AKRs1-5 tem um Loop-B curto, GDHs um Loop-B de tamanho médio e AKR7/AKR11 um Loop-B longo. O Loop-C é muito variável entre todas as sequências. Como a maioria dos AKRs, os GDHs têm duas  $\alpha$ -hélices externas ao barril, a hélice H1 dentro do Loop-B e a hélice H2 entre  $\alpha$ 8 e Loop-C (ambas as hélices são conservadas na maioria dos AKRs). A falta de uma descrição da estrutura 3D para GDH até agora impediu sua classificação adequada dentro da superfamília AKR.

## 3.2.5 Ligação ao NAD<sup>+</sup>

Os AKRs são caracterizados pela ligação de coenzimas de nucleotídeos de piridina, comumente NAD(P). Da mesma forma, apresentam um mecanismo de oxidação-redução conservado usando uma tétrade catalítica composta por aminoácidos altamente conservados (Asp, Lys, Tyr e His), com especificidade para substratos de açúcar ou esteróides <sup>149</sup> (Figura 26). Está bem estabelecido que GDH tem uma especificidade para NAD<sup>+</sup> como seu cofator (ao invés do NADP<sup>+</sup> mais comum) e que o açúcar L-gal é seu substrato preferido. <sup>61</sup> Para explorar completamente as diferenças entre GDH e outros membros da superfamília de proteínas AKR, também foi necessário resolver a estrutura de *So*GDH na forma de seu complexo com NAD<sup>+</sup>. Esta estrutura será referida aqui como *So*GDH-NAD<sup>+</sup>. A forma cristalográfica de *So*GDH-NAD<sup>+</sup> tinha duas moléculas na unidade assimétrica. O mapa *polder* para dos cofatores NAD<sup>+</sup> é dado na Figura 28 onde a qualidade da densidade eletrônica é claramente visível para toda a molécula.



Figura 26 - Alinhamento de múltiplo das sequências entre GDHs e AKRs. Cinco sequências de GDH de diferentes espécies (Uniprot: AB160990.1, OK632632.1, CAA20580.1, AK102223, AY176585) foram alinhadas; com sete representantes de diferentes famílias de AKRs (PDB:1AFS, 1YE4, 3UP8, 5JH1, 1MZR, 1GVE, 1PYF). Usando a sequência SoGDH do espinafre, a posição dos elementos da estrutura secundária (fitas em amarelo e hélices em vermelho) são indicadas. Da mesma forma, a numeração adotada para a descrição das posições dos aminoácidos foi baseada em SoGDH. Em azul escuro destacam-se os aminoácidos da tétrade catalítica, em azul claro os responsáveis pela ligação ao NAD<sup>+</sup>, em verde claro os aminoácidos importantes na determinação da especificidade do substrato e em roxo uma posição conservada nas ARKs que confere preferência ao cofator. NAD<sup>+</sup> ou NADP<sup>+</sup>). As posições com uma bola vermelha representam resíduos que interagem com NAD<sup>+</sup>

Fonte : Adaptada de VARGAS et al. 132



Figura 27 - Rede de interações com o cofator NAD<sup>+</sup> da SoGDH. As características específicas são destacadas de acordo com o seguinte código de cores; em azul escuro os aminoácidos da tétrade catalítica; em azul claro posições conservadas em todos os AKRs que interagem com NAD<sup>+</sup>; em magenta aminoácidos importantes para determinar a especificidade do substrato e em ciano moléculas de água. As linhas tracejadas indicam ligações de hidrogênio (preto), interações hidrofóbicas (roxo) e ligações de hidrogênio do doador pi (verde). As figuras foram geradas com PyMol v2.05 (Schrödinger, LLC) e Discovery Studio Visualizer V21.1.0 (BIOVIA, Dassault Systèmes).

Fonte : Adaptada de VARGAS et al. 132



**Figura 28** - Mapa *Polder* (a 2,5 σ) mostrando a qualidade do mapa de densidade eletrônica da molécula NAD<sup>+</sup> da cadeia A da estrutura SoGDH-NAD<sup>+</sup>.

Fonte : Adaptada de VARGAS et al. 132

As análises do PISA <sup>133</sup> indicaram que não se espera que as duas moléculas da unidade assimétrica formam o homodímero sob condições fisiológicas e é meramente o resultado do empacotamento do cristal. Isto é consistente com os resultados SEC-MALS apresentados anteriormente. Além disso, a análise SEC da enzima sob condições ácidas (como usado para

cristalização) ou na presença de NAD<sup>+</sup> indicou que ela era monomérica em todas as condições investigadas.

Até a presente data, os únicos membros da superfamília AKR que são conhecidos por serem homodímeros são aflatoxina dialdeído redutase (AKR7A1), xilose redutase dependente de NAD(P)H (AKR2B5) e a proteína supressora de auxotrofia de tirosina. <sup>151–153</sup> Uma interface de dimerização semelhante usando hélices  $\alpha$ 5,  $\alpha$ 6, H2 e Loop-C é observada em todas essas estruturas (**Figura 29A**). Por outro lado, em *So*GDH-NAD<sup>+</sup> a interface de contato entre os dois monômeros da unidade assimétrica é completamente diferente e envolve o loop  $\beta$ 5- $\alpha$ 3, Loop-B e Loop-C da cadeia A e as hélices  $\alpha$ 5,  $\alpha$ 6 e loop  $\alpha$ 5- $\beta$ 8 da cadeia B (**Figura 29B**). Consequentemente, as duas moléculas na unidade assimétrica de *So*GDH-NAD<sup>+</sup> interagem meramente como o produto de contatos de cristal e não tem relevância fisiológica. Esta observação é reforçada pela estrutura apo de *So*GDH que cristaliza com um monômero na unidade assimétrica em um grupo espacial que não possui eixos de rotação de ordem 2 e, portanto, dímeros no cristal são impossíveis. A soma total das evidências apresentadas aqui indica que o espinafre *So*GDH é uma proteína monomérica pelo menos nas condições usadas no presente trabalho e não um homodímero como relatado por *MIEDA et al.* <sup>62</sup>



Figura 29 - Interface de dimerização em AKRs homodiméricos (representados por aflatoxina aldeído redutase) comparada com a estrutura de SoGDH-NAD<sup>+</sup>. (A) A aflatoxina aldeído redutase (AKR7A1) usa hélices α5, α6, H2 e loop-C para dimerização (vermelho). (B) Na estrutura do SoGDH-NAD<sup>+</sup> os dois monômeros da unidade assimétrica fazem contatos cristalinos através do loop β5-α3, loop-B e loop-C da cadeia A e as hélices α5, α6 e loop α5-β8 da cadeia B (verde). Código PDB:1GVE da aflatoxina aldeído redutase.

Fonte: Adaptada de VARGAS et al. 132

Observa-se que a *So*GDH tem a tétrade catalítica conservada observada em AKRs: Asp57, Tyr62, Lys90 e His127 (**Figura 30**). A tétrade também faz parte de um grupo de resíduos que interagem diretamente com o NAD<sup>+</sup> (**Figura 30**). Isso sugere fortemente que a GDH segue o mesmo mecanismo catalítico básico descrito para outros AKRs, mas com preferência por favorecer a reação de oxidação (desidrogenação) em vez de redução (veja abaixo). Isso leva à produção de L-galactono-1,4-lactona a partir de L-galactose com a redução concomitante de NAD<sup>+</sup> para NADH.



Figura 30 - O sítio ativo em SoGDH. Uma superposição entre o sítio ativo de SoGDH (verde) e 3-αhidroxiesteróide desidrogenase (representante da família AKR1, em turquesa) mostra a tétrade catalítica (Asp27, Tyr62, Lys90 e His127) conservada em GDH. Código PDB: 1AFS de 3-α-HSD.

Fonte: Adaptada de VARGAS et al. 132

Os AKRs em geral possuem um grupo de resíduos conservados responsáveis pela ligação de NAD(P), sendo os mais conservados Asp57, Ser160, Asn161 e Gln183<sup>149,154</sup> (**Figura 31A e B**). Destes resíduos, o *So*GDH retém apenas Asp57, enquanto na posição 160 possui um Thr, em 161 a Gly e a 183 a Leu. Essas alterações levam à perda de duas ligações de hidrogênio com o anel de nicotinamida, formando novas interações de Van-der Waals com Gly161 e Leu183 (**Figura 31A e B**). Destaca-se também a substituição do resíduo aromático na posição 212 por Ala em *So*GDH (**Figura 31A e B**). Esses aminoácidos aromáticos, em AKRs, ocupam uma posição na parte inferior do anel de nicotinamida como visto na **Figura 31A** sendo característica notável do sítio de ligação do cofator. <sup>149,154</sup> Dada a sua importância estrutural e funcional, a falta deste resíduo aromático na GDH é, à primeira vista, surpreendente. No entanto, em GDHs, a interação pi-stack é recuperada com Tyr185, um aminoácido conservado nessas enzimas (**Figura 31B e Figura 26**).



Figura 31 - Comparação entre as interações de ligação de NADP<sup>+</sup> em AKRs e ligação de NAD<sup>+</sup> em SoGDH. Para melhor compreensão, as interações foram divididas nos lados nicotinamida e adenina que são tratados separadamente. (A) e (C) mostram as interações feitas com NADP<sup>+</sup> em 3-α-HSD (identificador de PDB: 1AFS). No lado da nicotinamida (A), são observadas três ligações de hidrogênio entre Ser160, Asn161 e Gln183, que juntamente com o "pi-stacking" formado por Tyr212, que orientam corretamente o anel de nicotinamida do NADP<sup>+</sup> para catálise. Do lado da adenina (C), destaca-se a interação entre Arg279 (conservada em AKRs) e o fosfato P2B e confere especificidade para NADP<sup>+</sup>. No lado da nicotinamida de SoGDH (B) as ligações de hidrogênio são perdidas, assim como o "pi-stack" na posição 212. No entanto, a orientação do anel de nicotinamida é mantida pela interação hidrofóbica de Leu183 e "pi-stack" de Tyr185. Além disso, uma ligação de hidrogênio é mantida entre Thr160 e o anel de nicotinamida. No lado da adenina (D), observamos que a ausência de P2B permite a interação via moléculas de água entre o Asn275 e a ribose. Da mesma forma, Gln279 interage diretamente com a ribose. Uma interação hidrofóbica envolvendo Phe33 e duas novas ligações de hidrogênio entre Asn283 e a adenina aparecem em SoGDH. PN: fosfato do lado da nicotinamida do cofator, PA: fosfato do lado da adenina. Ao comparar (B) com (A) ou (D) com (C), os resíduos conservados são apresentados na cor turquesa enquanto as substituições são mostradas em verde.

Este parece ser um exemplo clássico de substituição compensatória. Uma ligação de hidrogênio entre Thr27 e o anel de ribose no lado da nicotinamida do cofator em AKRs é mantida via Ser27. Além disso, em *So*GDH, uma nova ligação de hidrogênio é observada entre Ser213 e o grupo fosfato no lado da nicotinamida do cofator (PN) (**Figura 31B**).

Mudanças na interação entre a enzima e o cofator também são evidentes no lado da adenina do NAD<sup>+</sup> onde Lys/Arg273 e Arg279 normalmente conferem especificidade para o cofator NADP<sup>+</sup> na maioria dos AKRs (**Figura 31C**). <sup>149</sup> Em *So*GDH, a posição 273 é substituída por um Gly e 279 por um Gln (**Figura 31D**). Em AKRs, Lys/Arg273 formam uma ligação de hidrogênio com PN (o fosfato no lado da nicotinamida do cofator). No entanto, em *So*GDH, Gly273 faz uma ligação de hidrogênio usando sua cadeia principal para o fosfato no lado da adenina (PA). No entanto, o mais importante parece ser a perda de Arg279 em GDHs (**Figura 26 e Figura 31D**). A Arg279 forma ligações de hidrogênio com a fração fosfato P2B na ribose do lado da adenina, uma interação específica para enzimas que empregam NADP<sup>+</sup> como cofator <sup>149</sup> (**Figura 31C**). Em *So*GDH, a falta de Arg279 significa que a enzima é incapaz de compensar a carga formal em P2B tornando GDHs específicos para NAD<sup>+</sup> em vez de NADP<sup>+</sup>. No entanto, a substituição por Gln ainda permite a interação direta com um dos grupos hidroxila ribose no lado da adenina (**Figura 31D**).

Em AKRs, o resíduo 275 é altamente variável (**Figura 26**), mas isso não influencia sua interação com P2B porque isso ocorre através de sua amina de cadeia principal <sup>149</sup> (**Figura 31C**). No entanto, em GDHs, o resíduo homólogo, Asn275, é altamente conservado e, devido à ausência de P2B, Asn275 pode interagir com a porção ribose através de duas moléculas de água (**Figura 31D**). Na mesma região de *So*GDH, uma nova interação é observada. O resíduo Phe33 faz um contato hidrofóbico e uma ligação de hidrogênio doador pi com a ribose no lado da adenina (**Figura 31D**). Além disso, Leu283, que faz contato hidrofóbico com a adenina em AKRs, é substituído por Asn em *So*GDH. Isso permite a formação de duas ligações de hidrogênio adicionais com o anel de adenina, estabilizando ainda mais o cofator (**Figura 31D**).

Na **Figura 31** fica claro que o lado da nicotinamida do sítio de ligação do cofator é muito mais conservado do que o lado da adenina. No caso do primeiro, isto é esperado dada a necessidade, em ambos os casos, da cavidade do sítio de ligação acomodar a porção de nicotinamida orientada para catálise produtiva. No entanto, no caso do lado da adenina, embora alguns aspectos do sítio de ligação sejam preservados ao comparar AKRs e GDHs, em geral a cavidade foi radicalmente alterada para acomodar NAD<sup>+</sup> em vez de NADP<sup>+</sup>. A preferência por NAD<sup>+</sup> como cofator para GDH provavelmente está relacionada à direção da reação que está sendo catalisada que, para gerar AsA como produto, deve estar na direção

oxidativa. Conforme apontado por Barski et al., 2008<sup>155</sup> isso é diferente da maioria dos AKRs que geralmente favorecem a redução. Essas enzimas utilizam o NADPH como cofator que é mais abundante no citoplasma do que a forma oxidada NADP<sup>+</sup>, favorecendo a reação de redução <sup>155</sup>. Para NADH/NAD<sup>+</sup> ocorre o inverso, sendo NAD<sup>+</sup> a espécie predominante que ao preferir NAD<sup>+</sup> como cofator, a GDH favorece a produção oxidativa de L-galactono-1,4-lactona a partir de L-galactose e, assim, a produção efetiva de AsA.

## 3.2.6 Especificidade do substrato

Ainda não conseguimos sucesso na obtenção de cristais de SoGDH ligados ao substrato natural (L-Gal), um análogo do substrato ou um inibidor. No entanto, comparações com estruturas cristalográficas de AKRs complexados com inibidores ainda são possíveis para investigar a base estrutural das diferenças na especificidade do substrato. Para tanto, utilizaremos como exemplo os AKRs 3-a-hidroxiesteróide desidrogenase (3-a-HSD) e aldose redutase humana (ADR), ambos ligados a inibidores (Figura 32). Os AKRs apresentam aminoácidos dentro da bolsa de ligação do substrato que são responsáveis pela diferenciação entre uma ampla gama de esteróides e substratos de acúcares <sup>149,154</sup> (cor magenta na Figura 27). A Figura 32 mostra que no sítio ativo das três enzimas, os resíduos da tétrade catalítica, Asp57, Lys90, Tyr62 e His127 são totalmente conservadas (Figura 32). O Trp228 de Loop-B é essencial para a estabilidade e é conservado nas três proteínas <sup>149,154</sup>. Além disso, Ala59 e Trp93 desempenham o mesmo papel na estabilização do substrato em 3-α-HSD e ADR (Figura 32A e B). No entanto, em SoGDH essas posições são ocupadas por Ser59 e Arg93, o que pode ser importante para a diferenciação do substrato, uma vez que eles têm o potencial de formar novas ligações de hidrogênio com os grupos hidroxila no substrato L-Gal (Figura 32C). Os aminoácidos nas posições 61 e 128 estão envolvidos na diferenciação entre açúcares e esteróides.<sup>149</sup> Na posição 61, as hidroxiesteróides desidrogenases (HSDs) possuem Leu ou Ile, enquanto as aldose redutases (ADRs) possuem Val. Diferenças nesta posição podem alterar a topologia do sítio de ligação para acomodar substratos de diferentes tamanhos. Além disso, na posição 128, os HSDs têm Phe enquanto os ADRs têm um Trp conservado. A presença do nitrogênio indol permite a discriminação entre o açúcar e o esteróide, controlando a acessibilidade do substrato ou sua orientação no sítio ativo <sup>149</sup> (Figura 32A e B). A SoGDH de espinafre mostra diferenças altamente significativas nessas duas posições, que são ocupadas por Tyr61 e Asp128 (Figura 32C). O Tyr61 poderia potencialmente realizar uma interação de empilhamento com o anel de açúcar ou ligações de hidrogênio com o substrato por meio de seu grupo hidroxila. A Asp128, por outro lado, tem o potencial de formar um par de ligações de hidrogênio com hidroxilas adjacentes em L-Gal auxiliando no estabelecimento da orientação correta do substrato enquanto discrimina entre diferentes açúcares. Além disso, as enzimas de ligação à galactose geralmente usam um aminoácido aromático (Trp, Phe ou Tyr) para orientar corretamente o substrato de galactose dentro do sítio ativo. <sup>156</sup> Os GDHs, incluindo *So*GDH, têm um Tyr185 conservado que os diferencia dos AKRs restantes (**Figura 32C**). Além de restaurar a interação com o anel de nicotinamida NAD<sup>+</sup>, essa tirosina também pode desempenhar um papel no estabelecimento da especificidade da enzima para a L-galactose e orientá-la corretamente na bolsa de ligação (**Figura 32C**).



Figura 32 - Especificidade do substrato em AKRs e SoGDH. (A) 3-α-HSD ligado à testosterona (amarelo), mostrando aminoácidos importantes para a ligação do substrato. Leu61 e Phe128 são específicos para AKRs que se ligam a esteróides. (B) Aldose redutase humana (ADR) ligada ao inibidor "tolrestat". Val61 e Trp128 (verde) são específicos para AKRs que se ligam a açúcares. (C) O sítio ativo vazio como observado no complexo SoGDH-NAD<sup>+</sup>. Apesar de a SoGDH ter um açúcar como substrato natural, ela apresenta uma Tyr na posição 61 e uma Asp na 128. Também apresenta diferenças nas posições 59 (Ser) e 93 (Arg). Tyr185 parece ser específico para GDH, uma vez que as proteínas de ligação à galactose usam um aminoácido aromático para orientar o ligante para o sítio ativo. 3-α-HSD Código PDB: 1AFS, código ADR PDB: 2FZD.

Fonte: Adaptada de VARGAS et al. 132

O loop-C é essencial para determinar a especificidade do substrato em ARKs.<sup>149</sup> No entanto, em ambas as formas de *So*GDH relatadas aqui, o loop-C não está orientada para o sítio ativo devido à ausência de substrato e até agora não conseguimos esclarecer seu papel no caso de GDHs. Em resumo, apesar da ausência de um complexo com o substrato ou um análogo de substrato, as estruturas aqui apresentadas fornecem a base para testar, por mutagênese sítio dirigida, aqueles resíduos que são mais prováveis de estarem envolvidos no reconhecimento e seletividade do substrato.

Apesar da conservação do enovelamento ( $\beta/\alpha$ )8, as diferenças significativas observadas nos sítios de ligação do cofator e do substrato parecem justificar a classificação de GDHs como uma nova família dentro da superfamília AKR. Isso é confirmado por uma análise filogenética de sequências representativas das famílias existentes retiradas do banco de dados PKR da Universidade da Pensilvânia (https://hosting.med.upenn.edu/akr/) que mostra GDHs agrupados (**Figura 33**).



Figura 33 - O dendrograma mostra as relações topológicas entre alguns membros dos AKRs e 5 representantes de GDHs (em negrito). A árvore de consenso de "bootstrap" inferida a partir de 1000 réplicas mostra GDHs agrupados em um grupo monofilético (ramos vermelhos), sugerindo a formação de um novo subgrupo de AKRs por GDHs.

Fonte: Adaptada de VARGAS et al. 132

### 3.2.7 Estado aberto e fechado

Os AKRs têm um mecanismo de abertura e fechamento dependente de NAD(P), que permite a formação de um túnel que mantém o NAD(P) ligado. Essa estrutura transitória é formada após a ligação do cofator por uma mudança conformacional no loop-B e no loop  $\beta$ 3- $\alpha$ 1<sup>154</sup>. As interações que permitem a formação de túneis são diferentes em HSDs e ADRs. As ADRs diferem pela presença de uma ponte salina envolvendo Asp227, Lys35 e Lys273 e interações hidrofóbicas feitas por Trp27 (Loop- $\beta$ 3- $\alpha$ 1) e Pro226-Trp228 (Loop-B). Devido à mutação de Asp227 por uma Ser, os HSDs não possuem a ponte salina que ajuda a estabilizar o estado fechado. Consequentemente, em HSDs, o túnel é estabilizado por diferentes interações formadas entre vários resíduos de Loop- $\beta$ 3- $\alpha$ 1 (Pro33, Glu34, Lys35) e Loop-B (Asp225, Lys226, Thr227 e Trp228).<sup>154</sup>

Na *So*GDH os estados abertos e fechados são observados em função da ligação NAD<sup>+</sup>. Uma sobreposição das duas estruturas produz um RMSD de 0,66 para 318 átomos de C $\alpha$  com as regiões de maior divergência sendo Loop- $\beta$ 3- $\alpha$ 1 e Loop-B (importante para a formação do túnel) (**Figura 34**). Apesar de *So*GDH catalisar uma reação envolvendo açúcar, como ADRs, ela não possui uma ponte salina no túnel porque na posição 273, a Gly substitui o Lys/Arg conservados em ARKs. A formação e estabilização do Túnel são devidos as interações hidrofóbicas entre Val32-Phe33 (Loop- $\beta$ 3- $\alpha$ 1) e Pro226-Trp228 (Loop-B), bem como uma ligação de hidrogênio entre o nitrogênio indol de Trp228 e cadeia principal de Val32 (**Figura 34C**). embora haja uma variação significativa ente AKRs e GDH nos detalhes das interações responsáveis pela formação do túnel, este parece ser um mecanismo comum para todas essas enzimas, indicando que a ligação de NAD<sup>+</sup> induz uma mudança conformacional para a enzima que é presumivelmente essencial para uma catálise competente.

#### 3.2.8 Comparação da estrutura GDH

As estruturas primárias de GDHs mostram altos níveis de identidade de sequência entre diferentes espécies de plantas (percentual média de identidade de 76%), preservando aminoácidos-chave para a ligação de NAD<sup>+</sup> e aqueles deduzidos da discussão acima como sendo importantes para a especificidade do substrato de L-galactose (**Figura 32**). Portanto, espera-se que eles tenham uma estrutura tridimensional bem conservada, independente da espécie. O banco de dados de proteínas RCSB atualmente tem disponíveis as estruturas tridimensionais do arroz *Os*GDH em seus estados apo (7EZI) e holo (ligado a NAD<sup>+</sup>, 7EZL),

mas não há descrição publicada dessas estruturas. Conforme mencionado, o arroz *Os*GDH tem o enovelamento ( $\beta/\alpha$ )8-barril altamente conservada sendo semelhante ao espinafre *So*GDH.



Figura 34 - Os estados de abertura e fechamento do SoGDH. (A) A forma apo de SoGDH (cor rosa) mostra o estado aberto, com pouco contato entre Loop-B e Loop-β3-α1. (B) A forma holo de SoGDH-NAD<sup>+</sup> (cor verde) mostra o estado fechado, formando um túnel gerado pelo contato feito entre Loop-B e Loop-β3-α1 (NAD<sup>+</sup> é mostrado em preto dentro do túnel na parte inferior). (C) Uma superposição entre as formas SoGDH apo e holo mostram um "close-up" das regiões Loop-B e Loop-β3-α1. A ligação de NAD<sup>+</sup> resulta em contatos hidrofóbicos entre Val32, Phe33, Pro226 e Trp228, bem como ligações de hidrogênio entre o nitrogênio indol de Trp228 e a cadeia principal de Val32.

Fonte : Adaptada de VARGAS et al. 132

A forma apo de ambas as proteínas têm um RMSD de 0,45 Å para átomos de C $\alpha$  e as holoestruturas de 0,46 Å. Diferenças significativas existem apenas nas regiões Loop- $\beta$ 3- $\alpha$ 1 e Loop- $\beta$ 4- $\alpha$ 2 (**Figura 35A**). A importância da região Loop- $\beta$ 3- $\alpha$ 1 para a formação de túneis e o fato de que as sequências de aminoácidos nessas regiões são quase idênticas em ambas as enzimas, fazem com que essas diferenças sejam interessantes e dignas de uma investigação adicional. À primeira vista, não se esperaria que sequências semelhantes adotassem conformações tão radicalmente diferentes.

Uma conformação diferente de Loop- $\beta$ 3- $\alpha$ 1 na enzima do arroz deixa Phe33, necessária para a formação do túnel, em uma orientação oposta à observada no espinafre *So*GDH. Além disso, a mudança de posição do Loop- $\beta$ 3- $\alpha$ 1 (~6 Å) em direção ao Loop-B induz a formação de um túnel mesmo na forma apo, algo normalmente não observado em AKRs ou em *So*GDH (**Figura 35**). No entanto, analisando o mapa de densidade eletrônica de *Os*GDH fica evidente que esta região pode ter sido mal interpretada. Além disso, parece ter uma consequência indireta para o Loop  $\beta$ 4- $\alpha$ 2, que também adota uma conformação radicalmente diferente da observada em *So*GDH.

Como até agora não conseguimos cristais com qualidade de difração de camu-camu (*Md*GDH), realizamos sua previsão de estrutura usando Alphafold2. <sup>127</sup> O modelo resultante para o camu-camu *Md*GDH apresenta uma estrutura muito semelhante à do *So*GDH com um RMSD de 0,56 Å para os átomos de C $\alpha$ . Em *Md*GDH, um estado aberto é observado com Loop- $\beta$ 3- $\alpha$ 1 e Loop- $\beta$ 4- $\alpha$ 2 nas mesmas posições vistas em *So*GDH de espinafre (**Figura 35B**), e é muito diferente do observado na estrutura do arroz *Os*GDH. Não há diferenças notáveis entre *So*GDH e *Md*GDH ao redor do substrato e sítios de ligação de NAD<sup>+</sup>, o que implica que eles deveriam apresentar propriedades cinéticas semelhantes como foi observado experimentalmente (**Tabela 5**). Apesar do fato de que as previsões do AlphaFold2 ainda estão sob escrutínio pela comunidade científica, nossos resultados sugerem fortemente que a conformação descrita aqui para os loops em apo *So*GDH está correta e espera-se que seja conservada em outros GDHs. Uma vez que esta região está envolvida na formação do túnel e, portanto, é de importância funcional, parece que a interpretação estrutural da região apo da *Os*GDH necessita ser revisada.



Figura 35 - Comparações da estrutura GDH. (A) Uma superposição entre as estruturas cristalinas do espinafre SoGDH de espinafre (verde) e da OsGDH de arroz (roxo) na forma apo. Apesar de ter um RMSD de 0,45 Å para os átomos de Cα, a orientação do Loop-β3-α1 e do Loop-β4-α2 são muito diferentes (com um deslocamento de aproximadamente 6Å). (B) Superposição entre a estrutura cristalografica da SoGDH de espinafre (verde) e o modelo Alphafold2 para camu-camu MdGDH (azul). Loop-β3α1 e Loop-β4-α2 mostram a mesma orientação. Arroz OsGDH Código PDB: 7EZI.

Fonte : Adaptada de VARGAS et al. 132



**Figura 36** - Mapa de densidade eletrônica (2Fo – Fc em 1 $\sigma$ ) de Loop- $\beta$ 3- $\alpha$ 1 e Loop- $\beta$ 4- $\alpha$ 2 em SoGDH e SoGDH NAD<sup>+</sup>. A densidade mostra claramente a orientação de Phe33 em ambos os estados.

Fonte: Adaptada de VARGAS et al. 132

### 3.3 Subcapítulo: L-galactono-1,4-lactone desidrogenase (GalDH)

#### 3.3.1 Purificação e determinação do estado oligomérico de MdGalDH

Conforme descrito na metodologia, a expressão heteróloga e purificação de camu-camu *Md*GalDH foi realizada através de duas etapas. A primeira, constituída por uma cromatografia de afinidade ao níquel, essencial para isolar a proteína de interesse de todas as outras expressas da própria maquinaria da bactéria *E. coli*. As bandas observadas no gel de SDS-PAGE (**Figura 37A**) mostram uma proteína solúvel, com uma banda próxima de 66,2 kDa (massa esperada de 61,48 kDa). Este resultado mostra que *Md*GalDH purifica com uma baixa quantidade de contaminantes como observado nas frações de eluição E2-E5 (**Figura 37A**).



Figura 37 - Expressão e purificação de MdGalDH. A) Cromatografia de Afinidade de MdGalDH, M:Marcador de massa molecular, AI: antes de induzir com IPTG, DI: depois de induzir com IPTG, P: pellet-fração Insolúvel, S: fração solúvel, NL: proteínas que não interagem com o níquel imobilizado na resina, L1: lavagem com tampão, E1-E5: eluições com 25, 50, 100, 250 e 500 mM de imidazol. Eletroforese em gel SDS-PAGE 12%. B) Cromatografia de Exclusão de Massa Molecular (SEC) de MdGalDH, com picos homogêneos.

Fonte: Elaborada pelo autor

Na segunda etapa foi realizado uma cromatografia de exclusão de massa molecular, onde as eluiçoes que apresentaram uma banda única na etapa 1, foram concentradas e injetadas numa coluna de HiLoad Superdex 200 16/60. O cromatograma da purificação (**Figura 37B**) apresenta um pico bem definido e a pureza foi analisada por gel de SDS-PAGE 12%.

Para dar continuidade às próximas etapas de caracterização, foi importante definir a homogeneidade e estado oligomérico da proteína *Md*GalDH recombinante. Portanto, para estimar o estado oligomérico, utilizamos a cromatografia de exclusão de massa molecular

acoplada aos perfis de espalhamento de luz multi-ângulo (SEC-MALS) como técnica absoluta para determinar a massa molecular de nossa proteína em solução, assim como homogeneidade dela. A estimativa da massa obtida, mostrou um pico homogêneo, o qual corresponde a uma molécula de 60,69 kDa. Esta massa estaria relacionada com um monômero de GalDH, uma vez que a massa molecular teórica é 61,48 kDa (**Figura 38**). Este resultado é coerente com relatos da literatura sobre as enzimas de pimenta (*Capsicum annuum L, Ca*GalDH), <sup>25</sup> couve-flor (*Brassica oleracea, Bo*GDH), <sup>78</sup> e *Arabidopsis thaliana (At*GDH). <sup>38,65,157</sup>



Figura 38 - Cromatograma de SEC-MALS de MdGalDH. As curvas correspondentes à mudança no índice de refração diferencial normalizado (vermelho), intensidade de luz espalhada normalizada a 90° (azul) e peso molecular calculado do pico correspondente (linha preta horizontal) são fornecidos. A MW determinada de MdGalDH (60,69 kDa) que corresponde ao MW teórico esperado para um monômero de GalDH.

Fonte: Elaborada pelo autor

### 3.3.2 Enovelamento e estabilidade térmica de MdGalDH

Avaliamos o estado de enovelamento da proteína mediante o monitoramento de seu espectro de CD, a cada 2°C, em função do aumento de temperatura usando um gradiente de 20°C a 80°C. A **Figura 39** apresenta os espectros de CD da proteína *Md*GalDH em temperaturas onde ocorreram mudanças no seu sinal. Na temperatura de 20°C, podemos observar mínimos negativos em 222 e 208 nm que representam espectros característicos de proteínas que contêm elementos de estrutura secundária do tipo  $\alpha$  e  $\beta$  com predominância do sinal referente às contribuições de hélices  $\alpha$ . Os espectros após a temperatura de 48°C mostram uma diferença significativa, que está relacionado com a perda de elementos de estrutura secundária. A curva de desnaturação térmica, onde o sinal de espectro de CD foi monitorado a 222 nm, mostra a T<sub>m</sub> obtida (51,5 ± 0,2) após um ajuste de *Boltzman* no sinal detectado (**Figura 39**). Estudos da estabilidade e T<sub>m</sub> de algunas aldonolactonas oxidoredutases que

ligam covalentemente a FAD <sup>158-159</sup> e, de um GalDH de *Arabidopsis thaliana* (*At*GalDH) foram relatados com faixas de temperatura muito parecidos, variando entre 45 a 60°C aproximadamente, mostrando a similaridade entre integrantes da superfamília em termos de estabilidade térmica. Estes valores também são coerentes com experimentos demonstrando a perda de atividade catalítica das enzimas em temperaturas acima de 45-50°C. <sup>158-163</sup>



**Figura 39** - Espectro de CD de *Md*GalDH. A) Os espectros de CD obtidos em diferentes temperaturas. B) Desnaturação térmica, T<sub>m</sub> de 51,5°C.

Fonte: Elaborada pelo autor.

### 3.3.3 Parâmetros cinéticos de MdGalDH

A enzima recombinante do camu camu *Md*GalDH, apresenta cinética tipo Michaelis-Menten usando como seu substrato L-galactono-1,4-lactona (L-galL) e excesso de Citocromo C (2  $\mu$ M) (**Tabela 7**), com valor de K<sub>m</sub> de 0,044 mM, constante catalítica (k<sub>cat</sub>) de **5,85** s<sup>-1</sup> e uma eficiência catalítica (k<sub>cat</sub>/K<sub>m</sub>) de **132,95** mM<sup>-1</sup>s<sup>-1</sup> e pH ótimo de 8,0 (**Figura 40 A e B**). O isômero L-gulono-1,4-lactona (L-gulL) também foi oxidado, mas menos eficientemente do que L-galL (**Tabela 7**), o que significa que *Md*GalDH não é estritamente específico para Lgalactono-1,4, lactona mas apresenta uma preferência.



Figura 40 - Atividade de MdGalDH recombinante. (A) Cinética de Michaelis-Menten da oxidação mediada por MdGalDH de L-galL. (B) Atividade de MdGalDH em função do pH. (C) Inibição de MdGalDH pelo próprio substrato (L-galL). (D) Cinética de Michaelis-Menten da oxidação mediada por MdGalDH de L-gulL. As atividades foram medidas como descritas na metodología. Os valores são apresentados como a média ± SD de três experimentos.

Fonte: Elaborada pelo autor

Os valores de pH ótimo relatados para AtGalDH <sup>69, 163</sup>, estão na faixa entre 8 e 9.5 (**Tabela** 7), e valores semelhantes foram relatados para *Bo*GalDH. <sup>8,5,78</sup> De acordo com relatos recentes sob a atividade da GalDH para espécies nativas de plantas <sup>69,72,78,164-165</sup> o valor de K<sub>m</sub> reportado, varia de 0,044 a 3.3 mM, isso sem levar em consideração os valores relatados para mutantes de *At*GalDH (**Tabela 7**). A variação é bastante significativa com *Md*GalDH e *Bo*GalDH respectivamente. Por outro lado, o valore obtido para a proteína *Md*GalDH em relação a taxa de renovação com o substrato L-galactono-1,4-lactona é bem mais baixa do que o tipo selvagem do *At*GalDH e a maioria de suas variantes (**Tabela 7**). Enquanto a sua eficiência catalítica (k<sub>cat</sub>/K<sub>m</sub>) *Md*GalDH apresentou um valor 6 vezes mais baixo que *At*GalDH, e as variantes L56I, L56F e L56H também foram mais eficazes do que *Md*GalDH. A atividade de *Md*GalDH com L-gulono-1,4-lactona (L-gulL) é detectada com uma taxa de oxidação baixa (**Figura 40D**), como visto também em GalDH de batata-doce e tabaco (parâmetros não fornecidos) <sup>164-165</sup> sugerindo que esta reação não seja significativa *in vivo*.

Vale lembrar que em camu-camu a enzima especifica L-gulono-1,4-lactona desidrogenase (GulDH) apresenta uma alta afinidade pelo seu substrato natural (L-gulL) com um K<sub>m</sub> de 2,37  $\mu$ M <sup>166</sup> sendo provavelmente a enzima mais relevante para a oxidação deste substrato.

Tabela 7 – Parâmetros cinéticos de estado estacionário para GalDH estudados até o momento. As constantes cinéticas aparentes foram determinadas a 25°C em tampão de ensaio (pH 8,0) com concentrações variadas de aldonolactona, L-galactono-1,4-lactona e L-gulono-1,4-lactona (300 μM e 100 mM, respectivamente) e uma concentração constante de citocromo *c* (50 μM). Os valores são a média ± SD de pelo menos dois experimentos. ND= não detectável.

	L-galactono-1,4-lactona (L-galL)				L-gulono-1,4-lactona (L-gulL)			
Enzimas	Km (mM)	k <sub>cat</sub> (s <sup>-1</sup> )	$\frac{k_{cat}/K_m}{(mM^{-1}s^{-1})}$	Ki (mM)	pH ótimo	Km (mM)	k <sub>cat</sub> (s <sup>-1</sup> )	$\frac{k_{cat}/K_m}{(mM^{-1}s^{-1})}$
<i>Md</i> GalDH <sup>a</sup>	0.044	5.85	132.95	38,95	8	12.58	0,779	0,062
AtGalDH <sup>b</sup>	0.17	134	770	16,40	8-9,5	13,1	4,0	0,31
L56I <sup>b</sup>	0.32	240	750	ND	8,8	ND	ND	ND
$L56H^{b}$	0.121	32	260	ND	8,8	ND	ND	ND
L56F <sup>b</sup>	0.56	126	230	ND	8,8	ND	ND	ND
L56C <sup>b</sup>	0.99	76	78	ND	8,8	ND	ND	ND
L56A <sup>b</sup>	1,7	45	26	ND	8,8	ND	ND	ND
E386D <sup>c</sup>	12.1	4.7	0,39	ND	8,8	18,7	10,4	0,56
E386A <sup>c</sup>	17.9	1.4	0.078	ND	8,8	12,4	0,22	0,02
R388K <sup>c</sup>	123	48.5	0,40	ND	8,8	ND	ND	ND
R388A <sup>c</sup>	113	0,7	0,006	ND	8,8	ND	ND	ND
$BoGalDH^d$	3.3	ND	ND	32,60	8-8,5	ND	ND	ND

<sup>a</sup>GalDH parâmetros de acordo com este estudo.

<sup>b</sup>GalDH parâmetros de acordo com LEFERINK et al. <sup>72</sup>

<sup>c</sup>GalDH parâmetros de acordo com LEFERINK *at al.* <sup>69</sup>

<sup>d</sup>GalDH parâmetros de acordo com ØSTERGAARD et al. <sup>78</sup>

Fonte: Elaborada pelo autor.

Estudos de inibição pelo próprio substrato (L-galL) foram relatados para AtGalDH e BoGalDH (Ki= 16.40 e 32.60, respectivamente). A MdGalDH apresentou o mesmo fenômeno com um Ki=38,95 (**Figura 40C**) calculado pelo programa Prism-GraphPad. Os resultados sugerem que camu-camu MdGDH tem uma tolerância ligeiramente maior à inibição por L-galL que do BoGDH e AtGDH. A inibição poderia ser uma limitante da produção de vitamina C da via D-manose/L-galactose, mas acredita-se que isso não seria o único limitante. Um estudo recente em GalDH de pimenta revelou que a GalDH pode ser modulada por óxido nítrico (NO) em níveis transcricionais e pós-transcricionais, onde a expressão constante do gene GalDH poderia permitir um "turnover" contínuo da enzima, que pode manter ou aumentar os níveis de ascorbato. <sup>25</sup> Em contrapartida, um estudo sobre o silenciamento por interferência de RNA de GalDH de tomate mostrou que o conteúdo total de ascorbato permaneceu inalterado em plantas silenciadas com GalDH. <sup>73</sup> Estes resultados, sugerem a

possibilidade de que nas plantas poder-se-ia ativar vias alternativas de produção de vitamina C. Neste sentido, a regulação da produção do AsA seria ainda mais complexa do que imaginado.

## 3.3.4 Estrutura geral de *Md*GalDH

Até a presente data, a única estrutura reportada na literatura para GalDH foi a de *Brassica oleracea* resolvida mediante CRYO-EM. <sup>38</sup> No entanto, a descrição da estrutura não foi feita neste trabalho talvez por ser apenas um componente de um grande complexo (o complexo mitocondrial I). Neste trabalho, abordamos várias tentativas de cristalização para *Md*GalDH, das quais as condições mais representativas que resultaram em cristais com qualidade de difração de raios-X estão ilustradas na (**Figura 41**) Dados de difração de boa qualidade foram coletados para *Md*GalDH cristalizada em PACT (condição C5) e a estrutura foi refinada até uma resolução de 2.1 Å. A qualidade do modelo se reflete nos valores de  $R_{work}=21,55\%$  e  $R_{free}=24,43\%$ , com 96.98% de aminoácidos nas regiões mais favorecidas no diagrama de Ramachandran. Um resumo dos parâmetros da coleta e as estatísticas de refinamento são mostrados na **Tabela 8**.



Figura 41 - Cristais da proteína do Kit PACT de *Md*GalDH. (A) Condição C5 (0,1 M PCTP pH 8.0, 25% w/v PEG 1500), (B) condição D5 (0.1 M MMT pH 8.0, 25% w/v PEG 1.500), (C) condição A5 (0.1 M SPG pH 8.0, 25% w/v PEG 1500), (D) condição F9 (0,2 M Tartarato de sódio e potássio tetrahidratado, 0.1 M Bis-Tris-propano pH 6.5, 2% w/v PEG 1500. As proteínas cristalizaram em uma faixa de 3-7 mg/mL. As condições A e B foram as únicas que difrataram.

Fonte: Elaborada pelo autor

	<i>Md</i> GalDH			
Fonte de X-ray	Sirius MANACA			
Detetor	Pilatus 2M			
Parâmetros da cela: a, b, c (Å)	64.16, 70.09, 89.94,			
Parâmetros da cela: $\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 90.00, 90.00			
Grupo especial	$P2_{1}22_{1}$			
Resolução (Å)	44.97 - 2.1 (2.175 - 2.1)			
λ (Å)	1.3236			
Multiplicidade	13.3 (13.7)			
$R_{\rm pim}$ (all I+ & I-) (%)	0.064 (0.682)			
CC(1/2)	0.995 (0.706)			
Completeza (%)	100 (100)			
Reflexões	364118 (30336)			
Reflexões únicas	27421 (2210)			
$< I/\sigma(I) >$	7.9 (1.6)			
<i>R</i> (%) *	21,55%			
$R_{\rm free}$ (%) *	24,43%			
Número de átomos: proteína	3672			
Número de átomos: Água	145			
Número de átomos: ligante	0			
$B(Å^2)$	35.55			
Erro de coordenada (basado no-ML)	0.28			
(Å)	0.20			
Erro de fase (°)	25.35			
Ramachandran favorecida (%)	96.98			
Ramachandran permitida (%)	2.80			
Pontuação de confrontos de todos os	3 31			
átomos	5.51			
Comprimento de ligações (RMSD)	0.002			
(Å)*	0.002			
Ângulo de ligações (RMSD) (°) *	0.47			

Tabela 8 - Coleta de dados de raios-X e estatísticas de refinamento de estrutura de MdGalDH.

Fonte: Elaborada pelo autor

Neste trabalho de pesquisa, inicialmente tivemos dificuldades em encontrar uma solução por substituição molecular usando a estrutura de *Bo*GalDH resolvida por CRYO-ME (PDB:7A24), <sup>38</sup> enzima homóloga com 82.1% de identidade sequencial. A orientação dos "loops", a baixa resolução da estrutura de *Bo*GalDH (3.8 Å) e o valor de RMSD de 3.135 entre as duas estruturas (calculado após a determinação bem-sucedida de *Md*GalDH), podem ter sido os fatores principais do fracasso. No entanto, a estrutura da enzima *Md*GalDH foi finalmente resolvida por substituição molecular usando a estrutura de *Md*GalDH predita pelo

AlphaFold como modelo para a substituição molecular, usando o programa *Phaser* no pacote *Phenix*. <sup>118</sup>

A proteína *Md*GalDH faz parte das famílias de Flavoproteínas vanilil-álcohol oxidase (VAO), cujos membros se caracterizam por compartilhar um domínio de ligação de FAD conservado e um domínio de ligação ao substrato variável, que está posicionado sobre o anel de isoaloxazina do cofator. <sup>71,167</sup> A arquitetura variável da ligação do substrato permite grande variação na estrutura dos sítios ativos e na atividade catalítica das enzimas, mas mantendo o enovelamento de ligação ao cofator conservado. <sup>167</sup> A maioria das flavoenzimas são oxidorredutases, usam as propriedades de transferência de elétrons do sistema de anel isoaloxazina para catalisar a oxidação ou redução de dois elétrons de seus substratos. <sup>167</sup> Um estudo recente classificou a família VAO em 11 subgrupos com base na função e similaridade da sequência para uma melhor compreensão mecanicista das enzimas. <sup>167</sup> Neste contexto, *Md*GalDH faz parte do subgrupo 9: aldonolactona oxidorredutases cujos membros catalisam a etapa final da síntese de AsA ou compostos antioxidantes relacionados. <sup>167</sup>

Aldolactona oxidorredutases foram isoladas de várias fontes (plantas, animais e fungos), mas ainda não são muito bem caracterizadas. <sup>69</sup> Recentemente, a primeira estrutura tridimensional de uma aldonolactona oxidorredutase foi relatada <sup>38</sup>, mas a descrição geral, a natureza do sítio ativo e o mecanismo de reação não foram reportados. Neste trabalho, descrevemos pela primeira vez a estrutura cristalográfica de *Md*GalDH que apresentou uma molécula na unidade assimétrica (o monômero fisiológico) e a sua comparação com seu homólogo *Bo*GalDH. A estrutura de *Md*GalDH é dividida em dois domínios, o domínio de ligação ao FAD o qual possui 3 fitas  $\beta$  paralelas ( $\beta$ 1,  $\beta$ 2,  $\beta$ 3) e 5 fitas  $\beta$  antiparalelas ( $\beta$ 4,  $\beta$ 5,  $\beta$ 6,  $\beta$ 7 e  $\beta$ 8) flanqueadas por 5 hélices  $\alpha$  ( $\alpha$ A,  $\alpha$ B,  $\alpha$ C,  $\alpha$ D,  $\alpha$ E) na sua região N- terminal e 3 hélices  $\alpha$  ( $\alpha$ N,  $\alpha$ O e  $\alpha$ P) na sua região C-terminal. O N-terminal da cadeia e as regiões (resíduos: 58-63 e 244-247) são desestruturadas e não apresentam densidade eletrônica. O segundo domínio (o domínio de ligação ao substrato ou domínio "Cap") possui 7 fitas  $\beta$ antiparalelas ( $\beta$ 9,  $\beta$ 10,  $\beta$ 11,  $\beta$ 12,  $\beta$ 13,  $\beta$ 14 e  $\beta$ 15) flanqueadas por 7 hélices  $\alpha$  ( $\alpha$ F,  $\alpha$ G,  $\alpha$ H,  $\alpha$ I,  $\alpha$ J,  $\alpha$ K,  $\alpha$ L,  $\alpha$ M) (**Figura 42**).



**Figura 42** - Topologia de *Md*GalDH. (A) Representação em modelo de fitas do monômero de *Md*GalDH (esquerda), mostrando o enovelamento α/β característico da família VAO. O domínio de ligação ao FAD e o domínio de ligação ao açúcar são destacados em verde e turquesa respectivamente. As faixas em vermelho são os resíduos que participam direta ou indiretamente do sítio ativo. (B) Diagrama da topologia de *Md*GalDH, mostrando o domínio de ligação de FAD na cor verde com as 8 β-fitas antiparalelas (β1, β2, β3, β4, β5, β6, β7 e β8) e as 8 α- hélices (αA, αB, αC, αD, αE, αN, αO e αP). O domínio cap (turquesa) apresenta 7 β-fitas antiparalelas (β9, β10, β11, β12, β13, β14 e β15) e as 7 α-hélices (αF, αG, αH, αI, αJ, αK, αL, αM).

Fonte: Elaborada pelo autor

### 3.3.5 Sítio de ligação do cofator FAD

Ainda não conseguimos sucesso na obtenção de cristais de *Md*GalDH ligados ao cofator FAD. No entanto, realizamos uma busca no Protein Data Bank com o programa DALI <sup>168</sup>, para identificar estruturas semelhantes que tivesse FAD ligado. de possíveis resíduos que poderiam ligar ao FAD. O resultado mostrou que o homólogo mais próximo com estrutura cristalográfica em complexo com FAD foi a enzima Alditol oxidase (AldO) de *Streptomyces coelicolor* (PDB:2VFS). A enzima AldO está ligada ao FAD covalentemente no resíduo de His46, <sup>169</sup> é uma enzima monomérica de 418 aminoácidos e, pertencente à família das flavoproteínas VAO, especificamente do subgrupo 10: Alditol oxidases. <sup>167</sup> Todas as alditol oxidases caracterizadas contêm um cofator FAD ligado covalentemente e catalisam a oxidação dos grupos hidroxila primários de alditóis (álcoois de açúcar), produzindo aldoses, com oxigênio molecular atuando como aceptor de elétrons. <sup>167</sup> Embora GalDH se ligue ao FAD de modo não covalente, nós realizamos um alinhamento das sequências de aminoácidos

de ScAldO<sup>169</sup> comparando com outras 16 sequências aldonolactona oxidorredutase de outras espécies de plantas que tinham uma faixa de similaridade entre 47,9 - 96.2% (Figura 43). A similaridade de sequência de MdGalDH e ScAldO foi de 34.2%. Na Figura 43 se observa uma conservação alta dos resíduos que interagem direta e indiretamente com FAD na estrutura da AldO comparado com as outras 16 GalDHs. O FAD está profundamente enterrado na estrutura de AldO, onde seu grupo pirofosfato está envolvido em vários contatos de pontes de hidrogênio com átomos da cadeia principal. Os resíduos que interagem (Gly43 a Phe48) são conservados em todas as GalDHs, formando um motivo do grupo pirofasfato de FAD para as GalDHs, que estaria formado pelos resíduos "GxxLSP". Neste motivo a H46 (que liga covalentemente ao FAD) e F48 de ScAldO<sup>169</sup> são trocados por L56 e P58 em MdGalDH. A sobreposição de MdGalDH com ScAldO, mostra que os resíduos de ligação a FAD de ScAldO, sombreados com a cor turquesa (Figura 43), estão localizados na mesma posição e na mesma orientação em MdGalDH. Estudo de mutagênese sítio dirigida em AtGalDH revelaram que a presença de uma His na posição da L56 não resulta em uma ligação covalente do cofator. Aliás esta variante (L56H) mostrou uma eficiência catalítica 3 vezes menor que a nativa, o que sugere que o FAD foi mais fracamente ligado no mutante.<sup>160</sup> Todas as demais variantes de L56 estudadas podem ser facilmente reconstituídas pela adição de FAD.<sup>160</sup> Isso sugere que o acoplamento do cofator FAD talvez seja um processo autocatalítico, com a exigência de um sítio de ligação pré-organizado.<sup>70</sup>

Um estudo sugere que a importância de ter flavinilação covalente poderia estar no desenvolvimento da saturação do sítio ativo com o cofator para ambientes deficientes em flavina e o ancoragem do anel isoaloxazina do sítio ativo com as propriedades redox. <sup>72,170-171</sup> Os dados cristalográficos de *Sc*AldO, comparado com os de *Md*GalDH fornecem informações relevantes sobre os resíduos importantes da ligação ao FAD, permitindo uma visão geral da possível trajetória molecular da enzima com seu cofator e seu substrato, além de fornecer informações relevantes sobre os resíduos envolvidos na formação da cavidade catalítica (discutido em outra seção).

MdC-1DH (100%)	
MaGaiDH (100%)	MIAPP
RaGalDH (96.2%)	RYAPLPEDLHTVSNWSGTHEVOTRVFHOPESLAOLEEVVGEANAKKARIRPVGSGLSFNGIGLSRGMVNLALMDRVLEVDKERR
CsGalDH (83.7%)	RYAPLPEDLHTVSNWSGTHEVQTRVFHQPENIGELERVVKEANEKKTRIRPVGSGLSPNGIGLSRAGMVNLALMDKVLEVDKEKK
<i>Cs</i> GalDH*(81.8%)	RYAPLPEDLHTVSNHSGTHEVQTRNFHQPESVEELEKLVKEANEKRARI <mark>RPVGSGLSF</mark> NGIGLARAGMVNLALLDKVLEVDKEMK
AcGalDH (83.3%)	RYAPlpedlhtvsn <mark>a</mark> sgthevqtrtylqpetlqeleeivknanekkqki <mark>rpv<mark>gsglsp</mark>ngigltrqgmvnlalmdsvlevdkekk</mark>
VuGalDH (82.8%)	RYAP
AtGalDH (82.7%)	RYAPLPEDLHTVSNWSGTHEVQTRNFNQPERLADLEALVKESHEKKLRIRPVSSGLSPNGIGLSRSGWNLLALMKVLEVDKEKK
BoGalDH (82.1%)	
StGalDH (81.8%)	RYAPLPDDLHTVSNWSGTHEVRTRTFLQPESIEDLEGIVKEANVRKHKIRPVGSGLSPNGIGLTRAGMVNLALMDKVLSVDKEKK
SoGalDH (79.9%)	KYAPlpedlhtisn <mark>n</mark> sgthevhtrnfiqpesvsdleklvqeahqksqki <mark>rp</mark> v <mark>gsglsp</mark> ngvglsragmvnlglldkvlevdkekk
VaGalDH (77.6%)	REAAERDEDLHSVSN <mark>W</mark> SATHEVRTRILMQPESLDELERVVKDAHEKKQKIR <mark>P</mark> VGSGLSPNGIGLAQGGMVNLGLMNKVLEVDKDKK
GaGalDH (76.4%)	NYAPPQDLHSISNWSGTHEVQTRTFIQPESLSELEAVVEKCNGEKQKIRPVGSGLSPNGIGLTRSGMINLALMDRVLEVDEKTK
SvGalDH (77.1%)	KYAPLPEDLHAVSNWSATHEVHTRVLLQPDSLPALEDALATAHKERRKLRPFGSGLSPNGIGLSRAGWSLGLMDKVLDVDVKKK
KnGalDH (52.1%)	AVTOSKEPDVVTSVGDDDVETAMBSTASVATKALIQFBORELEKTVARADETGTKTEDVGGDADGAGGSSTEDAGDSGDAVGAVGAVGAVGAVGAVGAVGAVGAV
PrGalDH (47.9%)	EAEDHHEFVNWSATHECRPONFRVPETVDEVEKLLKKYHDKKOKLRCMGAGVSFNGLGFSGKTAGKSDANEAMMTLALLDRVLKVDKEKL
	$\mu_2$ $\alpha_B$ $\mu_{10}$ $\star$ $\alpha_{C}$ $\mu_{130}$ $\alpha_{D}$ $\mu_{0}$ $\mu_{150}$ $\mu_{1}$ $\alpha_{C}$ $\mu_{175}$ $\mu_{0}$
MdGalDH (100%)	TVRVEAGIRVQQLVDGIKEHGLTLQNFA IREQQIGGIVQVGAHGTGARLPPIDEQVISMKLVTPATGTIEVSKEKOPELFYLARGELGGUGVAEVTLQCVDRQ
SCALOU (34.2%)	TYKVGGGVKYARLAKVHARGLALPINA LIPHISVAGSVATOTHGSGVGNOSLASVVREVELVTADGSTVVLARD-EREGGATISLGALGVVTSLTLDLEPAT TYKVGGCVKYARLAKVHARGLALPINA LIPHISVAGSVATOTHGSGVGNOSLASVVREVELVTADGSTVVLARD-EREGGATISLGALGVTSLTLDLEPAT TYKVGGCVKYARLAKVHARGLALPINA LIPHISVAGSVATOTHGSGVGNOSLASVVREVELVTADGSTVVLARD-EREGGATISLGALGVTSLTLDLEPAT
CsGalDH (83.7%)	RVRVQAGIRVQELVDGIKDYGITUQIRASIREQQIGGIIQVGAHCTGARLPFIDEOIVSMKUVTPAKGIIPJSARDFELFYLARCGLGCLGCUVVAEVTICCVRRL
CsGalDH*(81.8%)	RVRVQAGIRVQGLVDEIKQYGLTLQNFABIREQQIGGIVQVGAHGTGAKLPPVDEQVISMKLVTPAKGTIEVSKEKDPDLFYLARCGLGGLGVVAEVTLQCVERQ
AcGalDH (83.3%)	RVRVQAGIRVQQLVDGIKDYGLTLQNFA <mark>S</mark> IREQQI <mark>GG</mark> IVQVGAHGTGARLPPIDEQVISMKLVTPAKGTIEISREKDPELFYLARCGL <mark>GGLGV</mark> VAEVTLQCVERQ
VuGalDH (82.8%)	TVRVQAGIRVQQLVDGLKDHGLTLQNFA <mark>S</mark> IREQQIGGIIQVGABGTGAKLPPIDEQVIAMKLVTPAKGTIEISKDKDPELFYLARCGL <mark>GGLGV</mark> VAEVTLQCVDRQ
AtGalDH (82.7%)	RVTVQAGIRVQQLVDAIKDYGLTLQNFATIREQQIGGIIQVGAHGTGARLPPIDEQVISMKLVTPAKGTIELSREKDPELFHLARCGLGGLGVVAEVTLQCVARH
NtGalDH (82.6%)	TVTVQAGIRVQQLVDAIKEYGITLQNFABIREQQIGGIVQVGAHGTGAKLPFIDEQVISMKLVTPAKGTIFISKEKDPELFYLARCCLGGIGVVAEVTLQCVERQ BUBVDACTEVDOVUDAIKEYGITLQNFABIREQQIGGIVQVGAHGTGAKLPFIDEQVISMKLVTPAKGTIFISKEKDPELFYLARCCLGGIGVVAEVTLQCVERQ
StGalDH (81.8%)	RVTVQAGTRVQQLVARLEGISLGILGUFALTEGISLGUFAGTGGUFAGTGARLPPTDEQVISMGUFATEISNEKDPELFVLARCGLGGLGVVARATIOVVERO
SoGalDH (79.9%)	TVRVEAGIRVQQLVDQIKEYGLTLQNFABIREQQIGGILQVGAHGTGARLPPIDEQVISVKLVTPAKGTIELSKEKDPELFYLARVGLGGLGVVAEVTLQCVDRQ
VaGalDH (77.6%)	TVRVQAGIRVQELVDEVKEYGLTLQNFA <mark>S</mark> IREQQIGGIIQVGAHGTGTRLPPIDEQIINMKLVTPSKGTIEVSKDKDPDLFYLARCGL <mark>GGLGV</mark> VAEVTIQCVDRQ
GaGalDH (76.4%)	RVRVEAGIRVQQLVDAIKEYGLTLENFA <mark>S</mark> IREQQIAGIVQVG <mark>AH</mark> GTGPKLPPIDEQVISLKLVTPAKGTIEVSKEKNPQLFHLARCGL <mark>G</mark> GL <mark>GV</mark> VAEVTLQCVDRH
SvGalDH (77.1%)	TVTVQAGIRVAELVDALREHGLTLQNFASIREQQVGGFIQVGAHGTGARLPPVDEQVISMKLVTPAKGTIELSREKDPELFYLARCGLGGLGVVAEVTLQCVERH
ESGAIDH (51.1%)	TVTVEAGARVQQVVDALAPHGITLQNYASISEQOLGGTQVGAHGTGATIPVDQVVGMKLVTPKLGTLELSNEKNPSLFRMAKVGLGAFGVVTEVTIKCVEKH OVEVDAGTVJOTASIA BEGIT I PENESISED POPOVOGEVTGAVI A VALGARVALVESSEK SEGANDET FEDI VEVGT GTGTVTTI TI OVEDBU
PrGalDH (47.9%)	VIVARIAVQULBALIBERGII LEMA I DOVVOGI VVOGI VVOGI VVOGVALKVI FARGI LI ISSOVIPELI RAVVIDAGALI VILLI VCVI O
. ,	
	$p_{20} \alpha F$ $p_{10} 225 p_{11} 248 \alpha G$
MdGalDH (100%)	ELVEHTTVSTIGEIKKNHKKFLSENKHVKYLYIPYDTVVVVTCNPVSKWKGPFKFKYKYTADEA-LQHVRQLYQESLQKYRP
RaGalDH (96.2%)	budged viele
CsGalDH (83.7%)	ELVEHTFISTMEDIKKNHKKLLSENKHVKYLYIPYTDAVVVTCNPVSKWKGAPKFNPKYTSDEA-LOHVRDIYKESLRKYRARD
CsGalDH*(81.8%)	elvehttvsnikeikknhkkllsenkhvkylhipytdtvvvvtcnpVSKwkGplkfkpkytkdea-lqhlrdlyreslkkyrygcylwytfcfykfgh
AcGalDH (83.3%)	ELVEHTFVSNMKEIKKTHKKLLSENKHVKYLYIPYTDTVVVVTCNPVSKWKGPPKFKSKYSHDEA-IQPVRDLYQESLKKYRGEE
VuGalDH (82.8%)	ELVEHTFVSSMKEIKRNHKKLLSDNKHVKYLYIPYDTVVVRCNPVSKWKGPPKFKPQYTKDEA-IQHVRDLYRESLQKYGAGG
N+CalDH (82.7%)	ELVEHTIVSNUGEIRRNHRRLLSANKHVRILTIPTDTVVVTCNPVSRWSGPRGRFRITTDEA-VQHVRDLIRESIVRIKVQD
BoGalDH (82.1%)	ELLEHTYVSTLEEIKKNHKKLLSTNKHVKYLYIPYTDTVVVTCNPVSKWSGAPKDKPKYTTEEA-LKHVRDLYRESIVKYRVOD
StGalDH (81.8%)	ELVEHTFLSNMKDIKKNHKKFLSENKHVKYLYIPYTDAVVVVTCNPMSKRKGPPKNKPKYTTEEA-LQHVRDLYLESLTKYRGQ
<i>So</i> GalDH (79.9%)	ELVEHTYISNMDEIKKNHKKLLSENKHVKYLYIPYTETVVVVTCNPVSRWSRPKNFKPKYTKDEA-LQHVRELYRNSLEKYR
VaGalDH (77.6%)	QLVEHTFVSNINEVKKNHKKFLSENKHVKYLHIPYTDSVVVVRCNPVSWWKGPPKFKPKYSRDEA-MQHVRDLYWDSLKNRP
GaGalDH (76.4%)	QUVENTIVSIMMEIKKNIKKLISENKHIKIIIYIPYTDAVVVICNPDSKWKGPPKNEPKFFEEEA-LQUVELIYKSLKKIK
EsGalDH (51.1%)	ULVESTAVMTRAEVKANHDALLRSNKHIRVMWIPYTDAVVVVRCNPVDLHGKSTTRPVOMDKALORGARDDN-TVDMTPANNERDKRDCLOPLEN
KnGalDH (52.1%)	KLVEKTWTVSRGEAKRNLGKWLREYRHIRYMWIPYTDTVVVVASNPEGSERGWFSWLWGAAKPAKQYSEDEQ-LAPARTLMREVEATKHPP
PrGalDH (47.9%)	KLIEKTTVMTLDEIRPNHNRWIVEFQHLRYMWIPYTDAVVVVQCRRAQ EDEPLNEKRFPPVQNSDKVRLEAPRKLYLELTKGKP
	$\alpha H = \alpha H = \alpha I $
MdGalDH (100%)	DVKSSNEDEPDINELSFTELRDKLLALDP-LNKDHVIKVNOAEAEFWKKSEGYRVGWSDEILGFDCGGOOWVSETCYPAGTLSKPSMKDIEYIEELK
ScAld0 (34.2%)	IDGFPYAAPAAEKMHPVPG <mark>M</mark> PAV <mark>N</mark> CTEQFGVPGPWHERLPHFRAE <mark>F</mark> T <mark>P</mark> SSG <mark>AE</mark> LOSEYLMPREHALAALHAMDAIR
RaGalDH (96.2%)	DVKSSNKDEPDIDELSFTELRDKLLALDP-LNKDHVIKVNQAEAEFWRKSEGYRVGWSDEILGFDCGGQQWVSETCFPAGTLSKPSMKDIEYIEELK
CsGalDH (83.7%)	ITAPSEGNVEQDINDLTFTELRDKLLALDP-LNKDHVIQVNQAEAEFWRKSEGYRVGWSDEILGFDCGGQQWVSETCFPVGTLSKPNMKDLQVIEDLK
CSGalDH* (81.8%)	KORADWITAKSPDGTEPDINELSFTELROKLLALDP-LNKEHVIKVNQARAEFWIKSEGYRVGMADEILGFDCGGQQWVSETCPPSGTLAKLSMDLETTEELK
VuGalDH (82.8%)	SRGESAE EXCIDELS FELEROLIALDO INKON V V VAVABABAE WILSBOILDE LOB COGYQUE DE LOF PAGILINES MADU V VAVABABAE VILSBOILDE LOB COGYQUE DE LOF PAGILINES MADU VILSBOILDE VI
AtGalDH (82.7%)	SGKKSPDSEPDIQELSFTELRDKLLALDP-INDVHVAKVNQAEAEFWKKSEGYRVCWSDEILGFDCGGQQWVSESCFPAGTLANPSMKDLEYIEELK
NtGalDH (82.6%)	vadsgsp-epeidelsftelrdklaldp-lnkehvieinkaeaefwrksegyrvgwsdeilgfdcgghowysetcfpagtlskpsmkdleyieelmuserseturgerset
BoGalDH (82.1%)	SKKTPDSREPDINELSFTELRDKLIALDP-LNDVHVGKVNQAEAEFWKKSEGYRVGWSDEILGFDCGGQQWVSETCFPAGTLAKPSMKDLEYIEQLK
StGalDH (81.8%)	VTDSGSPDEPEIVELSCTELROKLLAMDP-LNKEHVIKVNKAEAEVWRKSEGYRVCWSDEILGPDCGGHQWVSETCFPAGTLSKPSMKDLEYIEELM
SOGAIDH (79.9%)	vsrienipulsoffelkoklaid-ande-anabiilevnoakaefwkssejnvosdeilgfocggowysettepagtiakpnnkolevieks servenddamoisetteiddillaid-ande-inkuvikinoakaefwkssejnvosdamoti (genoggowysettepagtiakpnikuvi)
GaGalDH (76.4%)	KINSSMASEENKPLSFTELROHLIALDP-LINKHIVKINEAEAEFWIKKSEGIKVGWOBEILGFDCGGOWSETCFFGILBALDBWKDLEYVENLL
SvGalDH (77.1%)	TETESNDPEIDTLSFTELRDKLLALDP-LDKDHVVKVNKAEAEYWKKSEGYRMGWSDEILGFDCGGQQWVSENCFPTGTLAKPSTKDLDYIEKLL
EsGalDH (51.1%)	-LLLEKVGSTLSPAALKDISSLDFAGLRDHLLAVAP-LDVSWIKQVNNAEAEFWRRSQGTRVDYSDKILGFECGGQQWVNEVSFPVGTLEEPNGKDLDYMDELM
KnGalDH (52.1%)	Qapklppdal0sltffelrdkilaadp-ldwehvkkvneaeaefwkmregSrvawsheilQfdCgGQQw <mark>vleaafp</mark> agtvnkpSdasiafmeevl
PrGalDH (47.9%)	EPDYLDWSFTKLRDKLLAINP-LDKKHVIRVNEVEKLFWKLSEGYRVAYSDDIIGFDCGGQQL <mark>V</mark> EEVSFPTAGTLELDFMQKLL

continua

#### continuação

	×	N				
	$\alpha K$ (-B14)	<u> </u>		αL (	$\rightarrow \alpha N$	$\alpha N$
	375 ** 40	·· <del>[/1/</del>	425		450 ★	
MdGalDH (100%)	QLIEKEHIPAPAPIEQEWTARSKSPMSPASSPA	EDDIFSWVGI	IYLPTGDARQRKEITEEE	FHYRRMTQEWLWDR	YSAYEHWA <mark>K</mark>	IEVPKDNEELAALQARL
ScAld0 (34.2%)	ETLAPVLQTCEIRTVAADAQWLSPAYG	R-D-TVAAHFTW	VEDTAAVLPVVRRLEE#	LVP:	FAARPHWG <mark>K</mark>	VFTVPAG-ELRALYPRL
RaGalDH (96.2%)	RLIEKEDIPAPAPIEQEWTACSKSLMSPASSPA	EDDIFSWV <mark>G</mark> I <mark>I</mark> M	IYLPTSDARQRKEITEEE	FHYRRLTQERLWDR	YSAYEHWA <mark>k</mark>	IEVPKDNEELAALQARL
CsGalDH (83.7%)	KLIEKEKVPAPSPIEQRWSACSKSPMSPASSSS	DDEIFSWVGI <mark>I</mark> M	IYLPTTDARQRKEITEEE	FHYRHLTQSKLWGR	YSAYEHWA <mark>k</mark>	IEVLKDKDELAALQERL
<i>Cs</i> GalDH*(81.8%)	QLIEKEDIPAPAPIEQRWTARSQSVMSPAYSSV	QDDIFSWV <mark>G</mark> I <mark>I</mark> M	IYLPTMDARQRKEITDEE	FNYRHLSQKQLWDQ	YSAYEHWA <mark>k</mark>	IEVPKDKEELAALQARL
AcGalDH (83.3%)	QLIEKEMIPAPAPI <mark>P</mark> Q <mark>R</mark> WTARSKSPMSPASSAA	EEDIFSWV <mark>G</mark> I <mark>I</mark> M	YLPTMDPRQRKEITEKE	FHYRRLTQTQLWDQ	YSAYEHWA <mark>k</mark>	IEVPKDKDELAALQARL
VuGalDH (82.8%)	QLIENEEIPAPAPI <mark>E</mark> Q <mark>R</mark> WTASSRSPLSPASSPS	EDDIFSWV <mark>G</mark> I <mark>I</mark> M	IYLPTMDARQRKDITEEE	FHYRHLTQATLWDR	' SAFEHWA <mark>k</mark>	IEVPKDKEELAALQARL
AtGalDH (82.7%)	KLIEKEAIPAPAPI <mark>F</mark> QRWTARSKSPISPAFSTS	EDDIFSWV <mark>G</mark> I <mark>I</mark> M	IYLPTADPRORKDITDEE	FHYRHLTQKQLWDQ	FSAYEHWA <mark>k</mark>	IEIPKDKEELEALQARI
NtGalDH (82.6%)	QLIEKESVPAPAPI <mark>B</mark> Q <mark>R</mark> WTACSRSRMSPAYSSA	DDDIFSWV <mark>G</mark> I <mark>I</mark> M	IYLPTMDARQRRQITEEE	FHYRHMTQAQLWDH	YSAFEHWA <mark>K</mark>	IEVPKDKEELAALQERL
BoGalDH (82.1%)	ELIQKEAIPAPSPI <mark>E</mark> Q <mark>R</mark> WTGRSKSPMSPAFSTA	EEDIFSWVGI <mark>I</mark> M	IYLPTADPRQRKDITDEE	FHYRHLTQAKLWDQ	YSAYEHWA <mark>k</mark>	IEIPKDKEELEALQERL
StGalDH (81.8%)	QLIEKESVPAPAPI <mark>E</mark> Q <mark>R</mark> WTACSKGRMSPAYSSV	DDDIFSWVGI <mark>I</mark> M	YLPTMDARQRRQITEER	FHYRHMTQAQLWDH	Y SAFEHWA <mark>K</mark>	IEVPKDKEELTALQARL
<i>So</i> GalDH (79.9%)	QLIEKEGIPAPAPI <mark>E</mark> Q <mark>R</mark> WTARSKSPMSPASS-I	DDDIFSWV <mark>G</mark> I <mark>I</mark> M	YLPTSDPRQRKEITDEE	FHYRFLTQSRLWDL	YSAYEHWA <mark>K</mark>	IEIPKNKEELVDLQARL
VaGalDH (77.6%)	QLIEKEEIPAPAPI <mark>E</mark> Q <mark>R</mark> WTARSRSLMSPASSTK	EDAIFSWV <mark>G</mark> I <mark>I</mark> M	IYLPTMDAHQRQEITEAE	FNYREITQSQLWDQ	YSAIEHWA <mark>K</mark>	IEVPKDKEKLEALQERL
GaGalDH (76.4%)	ELIEKNDVPSPAPI <mark>E</mark> Q <mark>R</mark> WTGRSRSVMSPAYSSS	TDDIFSWVGI <mark>I</mark> M	YLPTMDPRGRKRITEEE	FHYRWLTQKQLWDD	FSAYEHWA <mark>K</mark>	VEIPKEKDELTWLQTRL
SvGalDH (77.1%)	QLIEKEDIPAPAPIEQRWTAHSKSPMSPASSSE	EDDVFSWVGI <mark>I</mark> M	YLPTSDARQRKDITEEE	FNYRSLAQT-LWDD	YSAYEHWA <mark>K</mark>	IEVPKDKDELAELQARL
ESGalDH (51.1%)	RLIEREGIPAPAPI <mark>E</mark> Q <mark>R</mark> WTSSSNATLSPAHDGAS	PETVFSWV <mark>G</mark> V <mark>I</mark> M	IYLPADDVDARKRVTEAE	FAYRDLCKVRLWDR	YDCAVHWA <mark>K</mark>	IEKPRTAREAEDTRARL
KnGalDH (52.1%)	ELVEGEKIAAPAPI <mark>E</mark> Q <mark>R</mark> WTKKSDSPMSPGYAPPGAF	KDSIHSWV <mark>G</mark> I <mark>I</mark> M	<b>YLPTMELVERMRIRSA</b>	FEEYKKKCEVAVWEK	YGAVEHWA <mark>K</mark>	VEIPEEKEDQKRLRERF
PrGalDH (47.9%)	KRIDDEDIPAHSPI <mark>E</mark> Q <mark>R</mark> WTARSTSLMSPASSSN	IPNQIFSWV <mark>G</mark> V <mark>I</mark> L	YLPTAEKKVREAIKDRE	MDF-YATYRDFMEP	FGATEHWA <mark>K</mark>	IEWSTDATERQKMRDRL
	$\alpha N_{\alpha} = \alpha \Omega$					
MdGalDH (100%)	REFEVENTAREL DENKTLSNNKLDKLEPSSDS	0				
Scaldo (34 2%)	ADFCALACALDPACKETNAEVBCULAC	·×				
RaGalDH (96 2%)	REFEVELANKARRELEPING	P				
CsGalDH (83.7%)	REFEVELATING	v				
CsGalDH*(81 8%)	REFEVOSYNKARKELOPNETLSNNMLEKLEPLSD	T				
AcGalDH (83.3%)	RKREPVDTYNKARSKLDPNRTLSNNMLEKLEPLSDK	T				
VuGalDH (82.8%)	RKREPVDSYNKARKDLDPNRILSNNMLEKLEAOSNT	'T				
AtGalDH (82.7%)	RKRFPVDAYNKARRELDPNRILSNNMVEKLFPVSTT	A				
NtGalDH (82.6%)	KKKLPVDAYNOARKELDPNRILSNNMLEKLFI					
BoGalDH (82.1%)	RKRFPVDAYNKARRELDPNRILSNNMVEKLFPVSKT	'A				
StGalDH (81.8%)	KKKFPVDAYNOARKELDPNRILSNNMLEKLFPSSEA	v				
SoGalDH (79.9%)	KKRFPVDSYNKARKELDPNRILSNNMLEKLFPLSED	VHOKIK				
VaGalDH (77.6%)	KKRFPVDLYNKARNELDPNRILSNEMLEKLFPLAET	'M				
GaGalDH (76.4%)	RKKFPVDEYNRARKELDPNNILSNSILGKLFPMHE-					
SvGalDH (77.1%)	RKRFPVDAYNKARMELDPNKVLSNAKLEKLFPALEF	VK				
EsGalDH (51.1%)	GRRFPVGAFNSVRALLDPKNIMANDHINSCLGQAK-					
KnGalDH (52.1%)	RRKYPLERYKEARRELDPKDLLGN-RLLKFISQEDK	VDES				
PrGalDH (47 9%)	NKRYPVDKFKKARDEVDPHHTLSNHTVDEMCA					

Figura 43 - Alinhamento de múltiplas sequências de Aldola de outras espécies de plantas e Alditol oxidase (ScAldO). As sequências usadas para o alinhamento foram: Myrciaria dubia (OK642676.1, MdGalDH), Streptomyces coelicolor (PDB:2VFR,ScAldO), Rhodamnia argentea (XP\_030517151.1, RaGalDH), Cannabis sativa (XP\_030495485.1, CsGalDH), Citrus sinensis (KAH9649046.1, CsGalDH\*), Actinidia chinensis var. chinensis (PSS01756.1, AdGalDH), Vigna umbellata (XP\_047172088.1, VuGalDH), Arabidopsis thaliana (BAA95212.1, AtGalDH), Nicotiana tabacum (NP\_001312485.1, NtGalDH), Brassica oleracea var. oleracea (PDB:7A23, BoGalDH), Solanum tuberosum (ADP09635.1, StGalDH), Spinacia oleracea (XP\_021838259.1, SoGalDH), Viscum álbum (UER43480.1, VaGalDH), Genlisea aurea (EPS69983.1, GaGalDH), Setaria viridis (XP\_034603601.1, SvGalDH), Ectocarpus siliculosus (CBJ29466.1, EsGalDH), Klebsormidium nitens (GAQ86319.1, KnGalDH), Phytophthora ramorum (KAH7509002.1, PrGalDH). O alinhamento foi realizado usando Clustal W.Os possíveis resíduos de ligação ao FAD (sombreado na cor turquesa) e substrato (sombreado na cor verde). O local de clivagem putativo da sequência de direcionamento mitocondrial na planta GALDH (R▼/YAP) é o início das sequências. A estrelha vermelha representa resíduos que formam a cavidade de ligação ao substrato. A His46 envolvido na ligação covalente do cofator FAD (cor azul, sombreado cor tusquesa).

Fonte: Elaborada pelo autor

Por outro lado, não comparamos com a estrutura de *Bo*GalDH por dois motivos: Primeiro a estrutura não apresenta o cofator ou substrato ligado e, segundo a estrutura apresenta uma orientação dos elementos de estruturas secundarias um pouco diferentes. A sobreposição das estruturas de *Md*GalDH e *Bo*GalDH (RMSD de 3.135), revela algumas diferenças. Por exemplo na estrutura de *Bo*GalDH a cauda N-terminal foi identificado ligada ao complexo I em maturação com os intermediários B14.5b e 15 kDa <sup>38</sup> (**Figura 44**). Precisamente esta região em *Md*GalDH não apresenta densidade eletrônica pelo fato dela não estar ligada ao

complexo, sugerindo que esta região é flexível, mas necessária para a formação do complexo I, quando assume uma estrutura definida. No entanto, a maior diferença é observada na região do domínio de ligação ao substrato. Em *Md*GalDH é notório a formação de mais estruturas secundarias ( $\alpha$ G e  $\beta$ 13) (**Figura 44**). Em *Bo*GalDH o sub-domínio formado pelo feixe de hélices- $\alpha$  ( $\alpha$ G,  $\alpha$ H e  $\alpha$ I em *Md*GalDH) é diferente por conta da falta de  $\alpha$ G e a presença de um loop mais comprido (**Figura 44C e D**). Por outro lado, o alinhamento mostra que não há diferenças significativas nas sequências de aminoácidos entre as duas proteínas nesta região (**Figura 43**). Portanto, isso sugere uma possível interpretação errônea da densidade eletrônica desta região em *Bo*GalDH. A resolução da estrutura de *Bo*GalDH (3.8Å) comparada com a de *Md*GalDH (2.1 Å) fortalece esta hipótese.



Figura 44 - Sobreposição da enzima MdGalDH e BoGalDH. A) Sobreposição das estruturas mostrando a diferença mais relevante no domínio de ligação ao substrato (círculo vermelho) B) Zoom e giro de 90° do domínio de ligação ao substrato mostrando as estruturas secundarias com maior diferença (círculo vermelho). C) Mostra a formação de uma hélice a mais em MdGalDH onde em BoGalDH se observa um "loop" muito flexível. D) Zoom da região das hélices de MdGalDH (3 hélices) x BoGalDH (2 hélices).

Fonte: Elaborada pelo autor

### 3.3.6 Mecanismo geral de MdGalDH

De forma geral, o mecanismo catalítico das aldonolactona oxidorredutase, é conhecido por ser semelhantes em todos os casos. <sup>167</sup> A ligação C2'-C3' do substrato lactona é oxidada a uma ligação dupla pela transferência de dois elétrons para o cofator flavina da enzima. A reoxidação subsequente do cofator flavina pode ocorrer por sua reação com o oxigênio molecular (L-gulono-1,4-lactona oxidase, GUO) ou aceptores de elétrons como o citocromo C (GalDH). <sup>167</sup> Mas, falando especificamente do mecanismo de GalDH para a produção de AsA, o mecanismo envolve duas semirreações, a redutiva e a oxidativa. Na etapa redutiva o cofator flavina oxidado é reduzido ao estado hidroquinona pelo substrato L-galactono-1,4-

lactona (L-galL). A enzima reduzida de dois elétrons é então reoxidada na semi-reação oxidativa pelo citocromo C. <sup>71,163</sup> Esta semi-reação oxidativa envolve duas etapas de um elétron e a formação transiente da flavina aniônica vermelha semiquinona. Este mecanismo foi estudado em *At*GalDH, usando a técnica de fluxo interrompido (*stop flow*). <sup>71,163</sup> A redução da flavina por L-galL parece ser o passo limitante da velocidade no ciclo catalítico. A reoxidação pelo Citocromo C (Cc) na semi-reação oxidativa ocorre relativamente rápido, com uma constante de velocidade bimolecular de  $3,6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ , quatro ordens de magnitude mais rápida que a reoxidação pelo oxigênio molecular. <sup>71,163</sup> Esta reoxidação do Cc revela a formação do complexo transitório de baixa afinidade com o Cc. Sua interação envolve movimentos dinâmicos proteína-proteína e fatores entrópicos. Além disso, revelaram que a afinidade de GalDH pelo Cc é semelhante para todos os diferentes estados redox. Esta interação relativamente não específica, não impede a rápida transferência de elétrons suficientes são aparentemente amostradas no complexo dinâmico. <sup>75</sup>

Até agora, não obtivemos sucesso na obtenção de cristais de *Md*GalDH ligado ao FAD e, nem na obtenção com o substrato ou inibidor. O fato de não se ter ainda a estrutura cristalográfica de aldonolactona oxidorredutase, complexada com o FAD, substrato ou inibidor dificulta nossa compreensão sobre os mecanismos que envolvem o sítio ativo da enzima *Md*GalDH.

No entanto, a elucidação de alguns resíduos importantes para a catálise foram relatados para *At*GalDH, usando mutagêneses sítio dirigida: C340, L56, Glu386 e Arg388, <sup>69,72,74</sup> mas a falta de uma estrutura tridimensional da enzima não permitiu desvendar como ocorre a interação do substrato no sítio ativo.

Nesse contexto, tentamos identificar e interpretar as ligações dos resíduos que estariam participando deste processo. Mas, devido ao limitante das estruturas cristalográficas da subfamília: Aldonolactonas oxidorredutases, achamos importante comparar a estrutura de *Md*GalDH com a estrutura de AldO, enzima homologa com maior similaridade estrutural. Embora o mecanismo catalítico de AldO seja diferente, os resíduos que participam da ligação do FAD e o substrato "xilitol" são conservados na sua maioria (**Figura 43**). Na estrutura de AldO a posição do xilitol é observada no fundo do funil do sítio ativo por meio de ligações de hidrogênio envolvendo seus grupos hidroxila com a Ser106 (átomos O2 e O3 do substrato), o Glu320 (O2 e O4) e a Thr345 (O5). Na estrutura de *Md*GalDH está conservado a Ser114,

Glu383 e, na posição da Thr345 é trocada por um resíduo de Ile412. Além disso, o átomo de O1 do xilitol estabelece interações de ligação de hidrogênio com dois grupos carregados positivamente: o átomo NZ de Lys375 e a porção guanidínio da Arg322. A presença do resíduo His343, localizado como "backdoor" da cavidade catalítica, define o tamanho e a forma da área acessível ao substrato em frente à flavina. Em *Md*GalDH é conservado Lys453, Arg385, mas a His343 é substituída por uma G410. Como resultado das interações feitas por estes resíduos, a cadeia alifática do xilitol em AldO fica "sanduíchado" entre o anel imidazol da cadeia lateral de His343 e o anel da flavina (**Figura 45A**). Este arranjo coloca a ligação C1-O1 do poliol bem na frente do átomo C4a do anel de flavina, em uma orientação para o evento de oxidação que pode ser tipicamente identificado por um mecanismo de transferência de hidreto. <sup>169</sup>

Por conseguinte, os resíduos de AldO (K375, T345, H343, E320, R322 e S106, (Figura **45**A) envolvidos na ativação do substrato, criam um aglomerado de doadores de ligações de hidrogênio que auxiliam na ligação e orientação do substrato para favorecer a desprotonação. A hipótese da posição de encaixe do L-galL na estrutura de MdGalDH foi realizada a partir uma sobreposição da estrutura AldO ligado com xilitol e MdGalDH. Para a sobreposição se levou em conta apenas os resíduos que interagem com o xilitol e a Figura 43 mostra os resíduos homólogos que estariam formando as possíveis interações em GalDH. Após a sobreposição e a análise dos resíduos que ligam com o ligante, a possível posição e orientação de L-galL, encaixada na estrutura de MdGalDH foi realizado manualmente no Coot (Figura 45B). A principal diferença observada entre a formação da cavidade do sítio ativo de AldO e MdGalDH é o resíduo de H343, que no caso da MdGalDH é trocado por uma G410. Precisamente a troca deste resíduo é estericamente relevante, pelo fato dela tornar a cavidade do sítio ativo de MdGalDH maior, para assim conseguir encaixar o substrato de acúcar de 6 carbonos de L-galL no sítio ativo (Figura 45C). Os demais resíduos do sítio ativo de MdGalDH são conservados nas aldonolactona oxidorredutases alinheadas (Figura 43) indicando semelhança no mecanismo catalítico. MdGalDH exemplifica de forma elegante como o processo evolutivo é econômico em gerar estruturas capazes de reconhecer substratos diferentes (portanto participando de vias metabólicas diferentes) enquanto simultaneamente preserve aspectos fundamentais dos processos químicos envolvidos.



**Figura 45** - Cavidade de ligação ao substrato xilitol em *Sc*AldO e L-galL em *Md*GalDH. A) Mostra a cavidade e os resíduos importantes para a catálise do xilitol de AldO. B) Mostra a cavidade e os possíveis resíduos que estariam participando da catálise de L-galL de *Md*GalDH. C) Sobreposição dos resíduos de *Md*GalDH e *Sc*AldO que formam a cavidade de entrada do substrato ao sítio ativo.

Fonte: Elaborada pelo autor

# 4 CONCLUSÃO

Nosso trabalho traz informações para preencher em parte a lacuna de conhecimento sobre os aspectos funcionais e estruturais das enzimas críticas envolvidas na biossíntese de vitamina C em plantas (**Figura 2**). Descrevemos parâmetros biofísicos, bioquímicos e estruturais de três enzimas da via D-manose/L-galactose de *Myrciaria dubia, camu-camu (Md*GME, *Md*GalDH, *Md*GDH) e da *Spinacia oloracea (So*GDH). Camu-camu apresenta uma elevada concentração de AsA, mas sua compreensão pela qual acumula esse AsA ainda é pouca compreendida, nós acreditamos que um dos motivos poderia estar envolvido com as funcionalidades das enzimas que participam da via.

Pela primeira vez as propriedades biofísicas e estruturais de MdGME recombinante foram descritas, os resultados obtidos fornecem uma melhor compreensão estrutural e funcional da enzima *Md*GME. Sua massa molecular para uma proteína dimerica de 94,22 kDa (Figura 6). A estrutura cristalográfica apresenta ser muito semelhante com a de AtGME, verificou-se que os resíduos catalíticos e da ligação ao NAD<sup>+</sup> são bem conservados, o alinhamento das sequências mostrou alto grau de conservação nos resíduos, sugerindo uma conservação estrutural de outras GMEs. No entanto, as estruturas complexadas de GME ligada só a NAD<sup>+</sup> e, GME complexada com NAD<sup>+</sup>-GDC/GDD (substrato e produto) nos revelaram uma melhor compreensão do mecanismo da enzima. Uma mudança conformacional da estrutura de GME é observado quando a enzima apresenta a ligação do NAD<sup>+</sup>, substrato e produto ao mesmo tempo, algo não descrito até o momento. A estrutura de *Md*GME-NAD<sup>+</sup> revelou 3 "loops" flexíveis que sofrem uma mudança conformacional na estrutura em função da presença do substrato/produto. Além disso, os mesmos loops estariam envolvidos na dimerização da enzima e do seu estado oligomérico funcional. Portanto, as informações fornecidas neste trabalho sobre GME é uma base fundamental para futuras pesquisas, envolvendo estudos de mutagênese sítio dirigida aos resíduos de importância catalítica e estrutural (dimerização) da enzima.

Para as GDHs recombinantes, descrevemos pela primeira vez sua estrutura cristalográfica em suas formas apo e holo, revelando assim sua base estrutural para a preferência por NAD<sup>+</sup> como cofator, importante para iniciar a reação na direção da síntese de AsA. A mudança conformacional observada entre as duas formas da enzima é semelhante à observada para a ligação do cofator das AKRs em geral, levando ao enterramento do NAD<sup>+</sup> dentro de um túnel de sítio de ligação com a nicotinamida direcionada para a região da tétrade catalítica. No entanto, apesar da ausência de uma estrutura cristalina da enzima com L-Gal ligada foi possível identificar resíduos que provavelmente estão envolvidos na seletividade do seu substrato natural. Estes incluem Tyr61, Tyr185, Ser59 e Asp128 e sua identificação abre novas perspectivas para experimentos futuros. No entanto, a semelhança das estruturas cristalográficas do *So*GDH com as AKRs, revela uma origem evolutiva comum, apesar da catálise ser diferente, redutiva para as maiorias das AKRs e de oxidação para as GDHs, o qual justifica o uso de NAD<sup>+</sup> como cofator ao invés de NADP<sup>+</sup>. Enquanto, a origem da seletividade da L-Gal permanece especulativa neste estágio devido à ausência de uma estrutura com um análogo de substrato ligado, o mecanismo regulatório por trás do controle da síntese de AsA em plantas é uma importante questão em aberto. De acordo com nossos resultados, tanto *Md*GDH quanto *So*GDH são refratários à inibição por AsA, indicando que em plantas de camu-camu e espinafre, a atividade catalítica de GDH não é regulada por inibição tipo "feedback". Relatos anteriores contrariando esta hipótese podem ter sido artefatos introduzidos por uma mudança no pH em vez de inibição competitiva.

A estrutura cristalográfica de uma aldonolactona oxidorredutase é descrita aqui pela primeira vez. No entanto, a ausência de uma estrutura cristalográfica ligada a FAD e/ou com L-galL/ou inibidor, fizeram com que comparássemos nossa estrutura cristalográfica com AldO (FAD, ligação covalente), homólogo mais próximo de GalDH com estrutura resolvida em complexo com FAD e substrato (xilitol). Graças a análise estrutural minuciosa e ao alinhamento da sequência de aminoácidos, pudemos identificar os possíveis resíduos que estariam interagindo com o sítio de ligação a FAD e L-galL (o substrato natural). Consequentemente, identificou-se os resíduos que estariam envolvidos na seletividade de encaixe do substrato natural (L-galL) a sítio ativo, entre eles a Ser114, Arg385, Lys453, Glu383, Ile411e finalmente a Gly409, resíduo que possivelmente estaria definindo o tamanho e forma da cavidade onde encaixa o substrato em frente à flavina. A identificação destes resíduos já abre novas perspectivas experimentais. Além disso, nossos resultados sugerem uma baixa eficiência catalítica de *Md*GalDH comparado com seu homólogo *At*GalDH e uma maior resistência a inibição pelo substrato.

Em suma, neste trabalho elucidamos detalhes estruturais das enzimas que participam das sínteses de AsA da via D-manose/L-galactose, estabelecendo as bases para futuros desenvolvimentos do uso destas proteínas para fins biotecnológicos. As observações das estruturas tridimensionais das três enzimas são importantes para o entendimento do mecanismo de catálise de cada uma delas. Os dados mostrados trazem novas informações, levantam hipóteses e confirmam alguns dados bioquímicos previamente estudados. A descrição de possíveis interações entre resíduos específicos e os substratos contribui para uma
compreensão mais ampla das relações estrutura-função das enzimas envolvidas na síntese da vitamina C.

## REFERÊNCIAS

1 SVIRBELY, J. L.; SZENT-GYÖRGYI, A. The chemical nature of vitamin C. *Biochemical Journal*, v. 26, n. 3, p. 865–870, 1932.

2 ENGLARD, S.; SEIFTER, S. The biochemical functions of ascorbic acid. *Annual Review* of Nutrition, v. 6, n. 1, p. 365–406, 1986.

3 PETERKOFSKY, B. Ascorbate requirement for hydroxylation and secretion of procollagen: relationship to inhibition of collagen synthesis in scurvy. *American Journal of Clinical Nutrition*, v. 54, n. 6 Suppl, p. 1135S-1140S, 1991.

4 PADAYATTY, S. J.; LEVINE, M. Vitamin c physiology: the known and the unknown and Goldilocks. *Oral Diseases*, v. 22, n. 6, p. 463–493, 2016.

5 STEINBERG, F. M.; RUCKER, R. B. *Padal. In*: LENNARZ, W. J.; LANE, M. D. (ed.). *Encyclopedia of biological chemistry.* 2nd ed. Waltham: Academic Press, 2013. p. 530–534.

6 BUETTNER, G. R. The pecking order of free radicals and antioxidants: lipid peroxidation, α-tocopherol, and ascorbate. *Archives of Biochemistry and Biophysics*, v. 300, n. 2, p. 535–543, 1993.

7 DAVEY, M. W. *et al.* Plant l-ascorbic acid: chemistry, function, metabolism, bioavailability and effects of processing. *Journal of the Science of Food and Agriculture*, v. 80, n. 7, p. 825–860, 2000.

8 LU, Y.; GUO, X. The effect of light in vitamin C metabolism regulation and accumulation in mung bean (*Vigna radiata*) germination. *Plant Foods for Human Nutrition*, v. 75, n. 1, p. 24–29, 2020.

9 SAMPAIO, I. *et al.* Modulation of beta-amyloid aggregation using ascorbic acid. *Biochimie*, v. 200, p. 36–43, 2022. https://doi.org/10.1016/j.biochi.2022.05.006

10 HANCOCK, R. D.; VIOLA, R. Biotechnological approaches for l-ascorbic acid production. *Trends in Biotechnology*, v. 20, n. 7, p. 299–305, 2002.

11 CHOTANI, G. *et al.* The commercial production of chemicals using pathway engineering. *Biochimica et Biophysica Acta (BBA)* - protein structure and molecular enzymology, protein engineering of enzymes, v. 1543, n. 2, p. 434–455, 2000.

12 BOUDRANT, J. Microbial processes for ascorbic acid biosynthesis: a review. *Enzyme and Microbial Technology*, v. 12, n. 5, p. 322–329, 1990.

13 MÆLAND, A.; WAAGBØ, R. Examination of the qualitative ability of some cold water marine teleosts to synthesise ascorbic acid. *Comparative Biochemistry and Physiology Part A:* molecular & integrative physiology, v. 121, n. 3, p. 249–255, 1998.

14 RUCKER, R. B.; SUTTIE, J. W.; MCCORMICK, D. B. Handbook of vitamins. Boca Raton: CRC Press, 2001.

15 HUANG, Z. *et al.* Glutathione enhances 2-keto-l-gulonic acid production based on Ketogulonicigenium vulgare model iWZ663. *Journal of Biotechnology*, v. 164, n. 4, p. 454–460, 2013.

16 ZHOU, M. *et al.* One-step biosynthesis of vitamin C in *Saccharomyces cerevisiae*. *Frontiers in Microbiology*, v. 12, p. 643472, 2021. DOI: 10.3389/fmicb.2021.643472.

17 ZOU, W. et al. Reconstruction and analysis of a genome-scale metabolic model of the vitamin C producing industrial strain Ketogulonicigenium vulgare WSH-001. *Journal of Biotechnology*, v. 161, n. 1, p. 42–48, 2012.

18 PAN, C.-H. *et al.* Reconstruction of amino acid biosynthetic pathways increases the productivity of 2-keto-l-gulonic acid in *Ketogulonicigenium vulgare-Bacillus endophyticus* consortium via genes screening. *Journal of Industrial Microbiology and Biotechnology*, v. 44, n. 7, p. 1031–1040, 2017.

19 DU, J. *et al.* Combinational expression of sorbose/sorbosone dehydrogenases and cofactor pyrroloquinoline quinone increases 2-keto-l-gulonic acid production in *Ketogulonigenium vulgare–Bacillus cereus* consortium. *Metabolic Engineering*, v. 19, p. 50–56, 2013. https://doi:10.1016/j.ymben.2013.05.006

20 CAI, L. *et al.* Genetic engineering of *Ketogulonigenium vulgare* for enhanced production of 2-keto-l-gulonic acid. *Journal of Biotechnology*, v. 157, n. 2, p. 320–325, 2012.

21 MANDLAA *et al.* Two-helper-strain co-culture system: a novel method for enhancement of 2-keto-l-gulonic acid production. *Biotechnology Letters*, v. 35, n. 11, p. 1853–1857, 2013.

22 JIA, N. *et al.* Comparative genomics analysis of the companion mechanisms of *Bacillus thuringiensis* Bc601 and *Bacillus endophyticus* Hbe603 in bacterial consortium. *Scientific Reports*, v. 6, p. 28794, 2016. https://doi: 10.1038/srep28794

23 CRAVENS, A.; PAYNE, J.; SMOLKE, C. D. Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nature Communications*, v. 10, p. 2142, 2019. https://doi.org/10.1038/s41467-019-09848-w

24 VICKERS, C. E.; BLANK, L. M.; KRÖMER, J. O. Grand challenge commentary: chassis cells for industrial biochemical production. *Nature Chemical Biology*, v. 6, n. 12, p. 875–877, 2010.

25 RODRÍGUEZ-RUIZ, M. *et al.* Characterization of the galactono-1,4-lactone dehydrogenase from pepper fruits and its modulation in the ascorbate biosynthesis. Role of nitric oxide. *Redox Biology*, v. 12, p. 171–181, 2017. https://doi.org/10.1016/j.redox.2017.02.009

26 RUNNING, J.; SEVERSON, D.; SCHNEIDER, K. Extracellular production of l-ascorbic acid by *Chlorella protothecoides*, Prototheca species, and mutants of *P. moriformis* during

aerobic culturing at low pH. *Journal of Industrial Microbiology and Biotechnology*, v. 29, n. 2, p. 93–98, 2002.

27 FENECH, M. *et al.* The role of GDP-L-galactose phosphorylase in the control of ascorbate biosynthesis. *Plant Physiology*, v. 185, n. 4, p. 1574–1594, 2021.

28 WHEELER, G. *et al.* Evolution of alternative biosynthetic pathways for vitamin c following plastid acquisition in photosynthetic eukaryotes. *eLife*, v. 4, p. e06369, 2015. https://doi.org/10.7554/eLife.06369

29 CARUSO, P. *et al.* A transcriptional analysis of the genes involved in the ascorbic acid pathways based on a comparison of the juice and leaves of navel and anthocyanin-rich sweet orange varieties. *Plants*, v. 10, n. 7, p. 1291, 2021.

30 CASTRO, J. C. *et al.* De novo assembly and functional annotation of *Myrciaria dubia* fruit transcriptome reveals multiple metabolic pathways for L-ascorbic acid biosynthesis. *BMC Genomics*, v. 16, p. 997, 2015. https://doi.org/10.1186/s12864-015-2225-6

31 LIAO, G. *et al.* Three metabolic pathways are responsible for the accumulation and maintenance of high AsA content in kiwifruit (*Actinidia eriantha*). *BMC Genomics*, v. 22, n. 1, p. 13, 2021.

32 VALPUESTA, V.; BOTELLA, M. A. biosynthesis of l-ascorbic acid in plants: new pathways for an old antioxidant. *Trends in Plant Science*, v. 9, n. 12, p. 573–577, 2004.

33 WHEELER, G. L.; JONES, M. A.; SMIRNOFF, N. The biosynthetic pathway of vitamin c in higher plants. *Nature*, v. 393, n. 6683, p. 365–369, 1998.

34 BULLEY, S.; LAING, W. The regulation of ascorbate biosynthesis. *Current Opinion in Plant Biology*, v. 33, p. 15–22, 2016.

35 LAING, W. A. *et al.* An Upstream open reading frame is essential for feedback regulation of ascorbate biosynthesis in arabidopsis. *Plant Cell*, v. 27, n. 3, p. 772–786, 2015.

36 CONKLIN, P. L. *et al. Arabidopsis thaliana* VTC4 encodes l-galactose-1-p phosphatase, a plant ascorbic acid biosynthetic enzyme. *Journal of Biological Chemistry*, v. 281, n. 23, p. 15662–15670, 2006.

37 LAING, W. A. *et al.* A highly specific l-galactose-1-phosphate phosphatase on the path to ascorbate biosynthesis. *Proceedings of the National Academy of Sciences*, v. 101, n. 48, p. 16976–16981, 2004b.

38 SOUFARI, H. *et al.* Specific features and assembly of the plant mitochondrial complex I revealed by cryo-EM. *Nature Communications*, v. 11, n. 1, p. 5195, 2020.

39 GOUDSMIT, E. M.; NEUFELD, E. F. Formation of GDP-L-galactose from GDP-Dmannose. *Biochemical and Biophysical Research Communications*, v. 26, n. 6, p. 730–735, 1967. 40 BARBER, G. A. Observations on the mechanism of the reversible epimerization of GDP-D-mannose to GDP-L-galactose by an enzyme from *Chlorella pyrenoidosa*. *Journal of Biological Chemistry*, v. 254, n. 16, p. 7600–7603, 1979.

41 BEERENS, K.; GEVAERT, O.; DESMET, T. GDP-mannose 3,5-epimerase: a view on structure, mechanism, and industrial potential. *Frontiers in Molecular Biosciences*, v. 8, 2022. https://doi.org/10.3389/fmolb.2021.784142

42 WOLUCKA, B. A.; DAVEY, M. W.; BOERJAN, W. A high-performance liquid chromatography radio method for determination of l-ascorbic acid and guanosine 5'-diphosphate-l-galactose, key metabolites of the plant vitamin c pathway. *Analytical Biochemistry*, v. 294, n. 2, p. 161–168, 2001.

43 WOLUCKA, B. A. *et al.* Partial purification and identification of GDP-mannose 3",5"epimerase of Arabidopsis thaliana, a key enzyme of the plant vitamin C pathway. *Proceedings of the National Academy of Sciences of the United States of America*, v. 98, n. 26, p. 14843– 14848, 2001.

44 WOLUCKA, B. A.; VAN MONTAGU, M. GDP-mannose 3',5'-epimerase forms GDP-Lgulose, a putative intermediate for the de novo biosynthesis of vitamin C in plants. *Journal of Biological Chemistry*, v. 278, n. 48, p. 47483–47490, 2003.

45 REITER, W. D.; VANZIN, G. F. Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant Molecular Biology*, v. 47, n. 1–2, p. 95–113, 2001.

46 TAO, J. *et al.* Molecular evolution of GDP-D-mannose epimerase (GME), a key gene in plant ascorbic acid biosynthesis. *Frontiers in Plant Science*, v. 9, 2018. https://doi.org/10.3389/fpls.2018.01293

47 LI, J. *et al*. Molecular cloning and functional analysis of a *Chrysanthemum vestitum* gme homolog that enhances drought tolerance in transgenic tobacco. *Scientific Reports*, v. 12, n. 1, p. 13551, 2022.

48 MOUNET-GILBERT, L. *et al.* Two tomato GDP-d-mannose epimerase isoforms involved in ascorbate biosynthesis play specific roles in cell wall biosynthesis and development. *Journal of Experimental Botany*, v. 67, n. 15, p. 4767–4777, 2016.

49 GILBERT, L. *et al.* GDP-d-mannose 3,5-epimerase (gme) plays a key role at the intersection of ascorbate and non-cellulosic cell-wall biosynthesis in tomato. *Plant Journal*, v. 60, n. 3, p. 499–508, 2009.

50 LI, X. *et al.* Biosynthetic gene pyramiding leads to ascorbate accumulation with enhanced oxidative stress tolerance in tomato. *International Journal of Molecular Sciences*, v. 20, n. 7, 2019.

51 MA, L. *et al.* Overexpression of an alfalfa GDP-mannose 3, 5-epimerase gene enhances acid, drought and salt tolerance in transgenic Arabidopsis by increasing ascorbate accumulation. *Biotechnology Letters*, v. 36, n. 11, p. 2331–2341, 2014.

52 IMAI, T. *et al.* Ectopic overexpression of peach GDP-d-mannose pyrophosphorylase and GDP-d-mannose-3',5'-epimerase in transgenic tobacco. *Plant Cell, Tissue and Organ Culture (PCTOC)*, v. 111, n. 1, p. 1–13, 2012.

53 BULLEY, S. M. *et al.* Gene expression studies in kiwifruit and gene over-expression in arabidopsis indicates that GDP-L-galactose guanyltransferase is a major control point of vitamin c biosynthesis. *Journal of Experimental Botany*, v. 60, n. 3, p. 765–778, 2009.

54 WOLUCKA, B. A.; VAN MONTAGU, M. The VTC2 cycle and the de novo biosynthesis pathways for vitamin c in plants: an opinion. *Phytochemistry*, v. 68, n. 21, p. 2602–2613, 2007.

55 MAJOR, L. L.; WOLUCKA, B. A.; NAISMITH, J. H. Structure and function of gdpmannose-3',5'-epimerase: an enzyme which performs three chemical reactions at the same active site. *Journal of the American Chemical Society*, v. 127, n. 51, p. 18309–18320, 2005.

56 WATANABE, K.; SUZUKI, K.; KITAMURA, S. Characterization of a GDP-d-mannose 3'',5''-epimerase from rice. *Phytochemistry*, v. 67, n. 4, p. 338–346, 2006.

57 VOXEUR, A. *et al.* Silencing of the gdp-d-mannose 3,5-epimerase affects the structure and cross-linking of the pectic polysaccharide rhamnogalacturonan II and plant growth in tomato. *Journal of Biological Chemistry*, v. 286, n. 10, p. 8014–8020, 2011.

58 LI, J. *et al.* Expression pattern and promoter analysis of the gene encoding gdp-d-mannose 3',5'-epimerase under abiotic stresses and applications of hormones by kiwifruit. *Scientia Horticulturae*, v. 150, p. 187–194, 2013. https://doi.org/10.1016/j.scienta.2012.11.008

59 GEVAERT, O. *et al.* Characterization of the first bacterial and thermostable gdp-mannose 3,5-epimerase. *International Journal of Molecular Sciences*, v. 20, n. 14, 2019. https://doi: 10.3390/ijms20143530

60 GEVAERT, O. *et al.* Gdp-altrose as novel product of GDP-mannose 3,5-epimerase: Revisiting its reaction mechanism. *International Journal of Biological Macromolecules*, v. 165, p. 1862–1868, 2020. https://doi.org/10.1016/j.ijbiomac.2020.10.067

61 GATZEK, S.; WHEELER, G. L.; SMIRNOFF, N. Antisense suppression of l-galactose dehydrogenase in *Arabidopsis thaliana* provides evidence for its role in ascorbate synthesis and reveals light modulated l-galactose synthesis. *Plant Journal*, v. 30, n. 5, p. 541–553, 2002.

62 MIEDA, T. *et al.* Feedback inhibition of spinach l-galactose dehydrogenase by l-ascorbate. *Plant and Cell Physiology*, v. 45, n. 9, p. 1271–1279, 2004.

63 LAING, W. A. *et al.* Kiwifruit l-galactose dehydrogenase: molecular, biochemical and physiological aspects of the enzyme. *Functional Plant Biology*, v. 31, n. 10, p. 1015–1025, 2004a.

64 MOMMA, M.; FUJIMOTO, Z. Expression, crystallization and preliminary X-ray analysis of rice l-galactose dehydrogenase. *Acta Crystallographica Section F:* structural biology and crystallization communications, v. 69, n. 7, p. 809–811, 2013.

65 SCHERTL, P. *et al.* L-galactono-1,4-lactone dehydrogenase (GLDH) forms part of three subcomplexes of mitochondrial complex I in *Arabidopsis thaliana*. *Journal of Biological Chemistry*, v. 287, n. 18, p. 14412–14419, 2012.

66 BARTOLI, C. G.; PASTORI, G. M.; FOYER, C. H. Ascorbate biosynthesis in mitochondria is linked to the electron transport chain between complexes III and IV. *Plant Physiology*, v. 123, n. 1, p. 335–344, 2000.

67 PINEAU, B. *et al.* L-galactono-1,4-lactone dehydrogenase is required for the accumulation of plant respiratory complex I. *Journal of Biological Chemistry*, v. 283, n. 47, p. 32500–32505, 2008.

68 BARTOLI, C. G. *et al.* Ascorbate content of wheat leaves is not determined by maximal l-galactono-1,4-lactone dehydrogenase (GalLDH) activity under drought stress. *Plant, Cell & Environment,* v. 28, n. 9, p. 1073–1081, 2005.

69 LEFERINK, N. G. H. *et al.* Functional assignment of Glu386 and Arg388 in the active site of 1-galactono-γ-lactone dehydrogenase. *FEBS Letters*, v. 583, n. 19, p. 3199–3203, 2009.

70 FRAAIJE, M. W. *et al.* A novel oxidoreductase family sharing a conserved FAD-binding domain. *Trends in Biochemical Sciences*, v. 23, n. 6, p. 206–207, 1998.

71 LEFERINK, N. G. H. *et al.* The growing VAO flavoprotein family. *Archives of Biochemistry and Biophysics*, Special Issue: enzymology in europe. v. 474, n. 2, p. 292–301, 2008b.

72 LEFERINK, N. G. H.; VAN DEN BERG, W. A. M.; VAN BERKEL, W. J. H. l-Galactonogamma-lactone dehydrogenase from Arabidopsis thaliana, a flavoprotein involved in vitamin C biosynthesis. *FEBS Journal*, v. 275, n. 4, p. 713–726, 2008a.

73 ALHAGDOW, M. *et al.* Silencing of the mitochondrial ascorbate synthesizing enzyme l-galactono-1,4-lactone dehydrogenase affects plant and fruit development in tomato. *Plant Physiology*, v. 145, n. 4, p. 1408–1422, 2007.

74 LEFERINK, N. G. H. *et al.* Galactonolactone dehydrogenase requires a redox-sensitive thiol for optimal production of vitamin C. *Plant Physiology*, v. 150, n. 2, p. 596–605, 2009a.

75 HERVÁS, M. *et al.* Communication between L–galactono–1,4–lactone dehydrogenase and cytochrome c. *FEBS Journal*, v. 280, n. 8, p. 1830–1840, 2013.

76 ZHANG, J. *et al.* A novel allele of l-galactono-1,4-lactone dehydrogenase is associated with enhanced drought tolerance through affecting stomatal aperture in common wheat. *Scientific Reports*, v. 6, p. 30177, 2016. https://doi.org/10.1038/srep30177

77 ÔBA *et al.* Purification and properties of l-galactono-  $\gamma$ -lactone dehydrogenase, a key enzyme for ascorbic acid biosynthesis, from sweet potato roots. *Journal of Biochemistry*, v. 117, n. 1, p. 120–124, 1995.

78 ØSTERGAARD, J. *et al.* Isolation of a CDNA coding forl-galactono-γ-lactone dehydrogenase, an enzyme involved in the biosynthesis of ascorbic acid in plants: purification, characterization, CDNA cloning, and expression in yeast. *Journal of Biological Chemistry*, v. 272, n. 48, p. 30009–30016, 1997.

79 CASTRO, J. C.; MADDOX, J. D.; IMÁN, S. A. Camu-camu—*Myrciaria dubia* (Kunth) McVaugh. *In*: RODRIGUES, S.; DE OLIVEIRA SILVA, E.; DE BRITO, E. S. (ed.). *Exotic Fruits*. New York: Academic Press, 2018. p. 97–105.

80 HERNÁNDEZ, M. S. *et al.* 16 - Camu-camu (*Myrciaria dubia* Kunth McVaugh). *In*: YAHIA, E. M. (ed.). *Postharvest biology and technology of tropical and subtropical fruits*.. Sawston: Woodhead Publishing, 2011. p. 352–375e. (Woodhead publishing series in food science, technology and nutrition)

81 JUSTI, K. C. *et al.* Nutritional composition and vitamin c stability in stored camu-camu (*Myrciaria dubia*) pulp. *Archivos Latinoamericanos de Nutrición*, v. 50, n. 4, p. 405–408, 2000.

82 ŠMÍD, J. *et al.* Morphological and genetic diversity of camu-camu [*Myrciaria dubia* (Kunth) McVaugh] in the Peruvian Amazon. *PLOS ONE*, v. 12, n. 6, p. e0179886, 2017.

83 IMÁN CORREA, S. *et al.* Contenido de vitamina c en frutos de camu camu *Myrciaria dubia* (H.B.K) Mc Vaugh, en cuatro estados de maduración, procedentes de la Colección de Germoplasma del INIA Loreto, Perú. *Scientia Agropecuaria*, v. 2, n. 3, 2011. https://doi.org/10.17268/sci.agropecu.2011.03.01

84 CASTRO, J. C. *et al.* Bioactive compounds of camu-camu (*Myrciaria dubia* (Kunth) McVaugh). *In*: MURTHY, H. N.; BAPAT, V. A. (ed.). *Bioactive compounds in underutilized fruits and nuts*. Cham: Springer International Publishing, 2019. p. 1–25. (Reference series in phytochemistry).

85 DO, N. Q. *et al.* Camu-camu fruit extract inhibits oxidative stress and inflammatory responses by regulating nfat and nrf2 signaling pathways in high glucose-induced human keratinocytes. *Molecules*, v. 26, n. 11, p. 3174, 2021.

86 ABOT, A. *et al.* Camu-camu reduces obesity and improves diabetic profiles of obese and diabetic mice: a dose-ranging study. *Metabolites*, v. 12, n. 4, p. 301, 2022.

87 LANGLEY, P. C. *et al.* Antioxidant and associated capacities of camu camu (myrciaria dubia): a systematic review. *Journal of Alternative and Complementary Medicine*, v. 21, n. 1, p. 8–14, 2015.

88 AZEVEDO, L. et al. Camu-camu (Myrciaria dubia) from commercial cultivation has higher levels of bioactive compounds than native cultivation (amazon forest) and presents

antimutagenic effects in vivo. *Journal of the Science of Food and Agriculture*, v. 99, n. 2, p. 624–631, 2019.

89 CORREA, S. I. *et al.* Contenido de vitamina c en frutos de camu camu *Myrciaria dubia* (H.B.K) Mc Vaugh, en cuatro estados de maduración, procedentes de la Colección de Germoplasma del INIA Loreto, Perú. *Scientia Agropecuaria*, v. 2, n. 3, p. 123–130, 2011.

90 LOCATO, V.; CIMINI, S.; DE GARA, L. Strategies to increase vitamin c in plants: from plant defense perspective to food biofortification. *Frontiers in Plant Science*, v. 4, 2013. https://doi: 10.3389/fpls.2013.00152

91 PINEDO PANDURO, M. *et al. Plan de mejoramiento genético de camu camu*. Iquito: Instituto de Investigaciones de la Amazonía Peruana, 2004.

92 CHIRINOS, R. *et al.* Antioxidant compounds and antioxidant capacity of peruvian camu camu (*Myrciaria dubia* (H.B.K.) McVaugh) fruit at different maturity stages. *Food Chemistry*, v. 120, n. 4, p. 1019–1024, 2010.

93 AGUIAR, J. P. L.; SOUZA, F. C. A. Camu-Camu (*Myrciaria dubia* HBK): yogurt processing, formulation, and sensory assessment. *American Journal of Analytical Chemistry*, v. 06, n. 05, p. 377, 2015.

94 GONÇALVES, J. N. *et al.* Elaboração de cerveja artesanal do estilo caxiri beer com adição de camu-camu (*Myrciaria dubia*). *Brazilian Journal of Science*, v. 1, n. 4, p. 101–108, 2022.

95 OLIVEIRA, R. F. P. *et al.* O desenvolvimento da biotecnologia industrial nos processos produtivos no estado do amazonas: the development of industrial biotechnology in production processes in the state of amazonas. *Brazilian Journal of Development*, v. 8, n. 8, p. 57836–57858, 2022.

96 MATOS, Í. T. S. R. *et al.* Yeasts with Fermentative Potential Associated with Fruits of Camu-Camu (*Myrciaria dubia*, Kunth) from North of Brazilian Amazon. *Scientific World Journal*, v. 2021, p. 9929059, 2021. DOI: 10.1155/2021/9929059.

97 LIM, T. K. *Myrciaria dubia. In*: LIM, T. K. (ed.). *Edible medicinal and non-medicinal plants:* volume 3, fruits. Dordrecht: Springer, 2012. p. 631–638

98 SALOMÃO-OLIVEIRA, A. *et al.* Ascorbic acid from lyophilized camu-camu fruit: stability and quality control of hard capsules. *Revista de Ciências Farmacêuticas Básica e Aplicada*, v. 37, n. 1, 2016. Disponível em: https://rcfba.fcfar.unesp.br/index.php/ojs/article/view/11

99 VARGAS, B. L. *et al.* Effect of camu-camu capsules on blood glucose and lipid profile of healthy adults. 2015. Disponível em: https://www.medigraphic.com/pdfs/revcubplamed/cpm-2015/cpm151e.pdf. Acesso em: 23 jan. 2021.

100 CASTRO, J. C. *et al.* Dataset of de novo assembly and functional annotation of the transcriptome during germination and initial growth of seedlings of *Myrciaria Dubia* "camucamu". *Data in Brief*, v. 31, p. 105834, 2020. https://doi.org/10.1016/j.dib.2020.105834

101 RAAB, D. *et al.* The geneoptimizer algorithm: using a sliding window approach to cope with the vast sequence space in multiparameter DNA sequence optimization. *Systems and Synthetic Biology*, v. 4, n. 3, p. 215, 2010.

102 FOLTA-STOGNIEW, E.; WILLIAMS, K. Determination of molecular masses of proteins in solution: Implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. *Journal of Biomolecular Techniques*, v. 10, n. 2, p. 51–63, 1999.

103 VELOURS, C. et al. Determination of the absolute molar mass of [Fe-S]-containing proteins using size exclusion chromatography-multi-angle light scattering (SEC-MALS). *Biomolecules*, v. 12, n. 2, p. 270, 2022.

104 SAHIN, E.; ROBERTS, C. J. Size-exclusion chromatography with multi-angle light scattering for elucidating protein aggregation mechanisms. *In*: VOYNOV, V.; CARAVELLA, J. A. (ed.). *Therapeutic proteins:* methods and protocols. methods in molecular biology. Totowa, NJ: Humana Press, 2012. p. 403–423.

105 WOODY, R. W. Theory of circular dichroism of proteins. *In*: FASMAN, G. D. (ed.). *Circular dichroism and the conformational analysis of biomolecules*. Boston: Springer, 1996. p. 25–67.

106 VENYAMINOV, S. Y.; YANG, J. T. Determination of protein secondary structure. *In*: FASMAN, G. D. (ed.). *Circular dichroism and the conformational analysis of biomolecules*. Boston: Springer, 1996. p. 69–107.

107 RANJBAR, B.; GILL, P. Circular dichroism techniques: biomolecular and nanostructural analyses- a review. *Chemical Biology & Drug Design*, v. 74, n. 2, p. 101–120, 2009.

108 LEES, J. G. *et al.* CDtool—an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Analytical Biochemistry*, v. 332, n. 2, p. 285–289, 2004.

109 MOBERLY, J. G.; BERNARDS, M. T.; WAYNANT, K. V. Key features and updates for Origin 2018. *Journal of Cheminformatics*, v. 10, n. 1, p. 5, 2018.

110 SOSSO, G. C. *et al.* Crystal nucleation in liquids: open questions and future challenges in molecular dynamics simulations. *Chemical Reviews*, v. 116, n. 12, p. 7078–7116, 2016.

111 ANWAR, J.; ZAHN, D. Uncovering molecular processes in crystal nucleation and growth by using molecular simulation. *Angewandte Chemie International Edition*, v. 50, n. 9, p. 1996–2013, 2011.

112 SEAR, R. Quantitative studies of crystal nucleation at constant supersaturation: experimental data and models. *CrystEngComm*, v. 16, n. 29, p. 6506–6522, 2014.

113 WINTER, G. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*, v. 43, n. 1, p. 186–190, 2010.

114 VAGIN, A.; LEBEDEV, A. MoRDa, an automatic molecular replacement pipeline. *Acta Crystallographica Section A*, v. 71, n. a1, p. s19–s19, 2015.

115 VAGIN, A.; TEPLYAKOV, A. MOLREP: an automated program for molecular replacement. *Journal of Applied Crystallography*, v. 30, n. 6, p. 1022–1025, 1997.

116 LANGER, G. *et al.* Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols*, v. 3, n. 7, p. 1171–1179, 2008.

117 MCCOY, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica Section D, Biological Crystallography*, v. 63, n. Pt 1, p. 32–41, 2007.

118 ADAMS, P. D. *et al.* Phenix: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallographica Section D, Biological Crystallography*, v. 66, n. Pt 2, p. 213–221, 2010.

119 EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D, Biological Crystallography*, v. 60, n. Pt 12 Pt 1, p. 2126–2132, 2004.

120 BRAMUCCI, E. *et al.* PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinformatics*, v. 13, n. 4, p. S2, 2012.

121 RAUF, M. A.; ZUBAIR, S.; AZHAR, A. Ligand docking and binding site analysis with pymol and autodock/vina. *International Journal of Basic and Applied Sciences*, v. 4, n. 2, p. 168–177, 2015.

122 GANESH, K. *et al.* Dynamic approach to predict pH profiles of biologically relevant buffers. *Biochemistry and Biophysics Reports*, v. 9, p. 121–127, 2017. https://doi.org/10.1016/j.bbrep.2016.11.017

123 THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, v. 22, n. 22, p. 4673–4680, 1994.

124 GOLDMAN, N. Maximum likelihood inference of phylogenetic trees, with special reference to a poisson process model of DNA Substitution and to parsimony analyses. *Systematic Biology*, v. 39, n. 4, p. 345–361, 1990.

125 FELSENSTEIN, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, v. 39, n. 4, p. 783–791, 1985.

126 TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, v. 38, n. 7, p. 3022–3027, 2021.

127 JUMPER, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature*, v. 596, n. 7873, p. 583–589, 2021.

128 MIRDITA, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods*, v. 19, n. 6, p. 679–682, 2022.

129 STEINEGGER, M.; SÖDING, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, v. 35, n. 11, p. 1026–1028, 2017.

130 LASKOWSKI, R. A. *et al.* Procheck: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, v. 26, n. 2, p. 283–291, 1993.

131 EISENBERG, D.; LÜTHY, R.; BOWIE, J. U. [20] Verify3d: assessment of protein models with three-dimensional profiles. *In: Methods in Enzymology Macromolecular Crystallography Part B*, v. 277, p. 396–404, 1997. https://doi.org/10.1016/S0076-6879(97)77022-8

132 VARGAS, J. A. *et al.* Structural characterization of l-galactose dehydrogenase: an essential enzyme for vitamin c biosynthesis. *Plant and Cell Physiology*, v. 63, n. 8, p. 1140–1155, 2022.

133 KRISSINEL, E.; HENRICK, K. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, v. 372, n. 3, p. 774–797, 2007.

134 ELEZ, K.; BONVIN, A. M. J. J.; VANGONE, A. Biological vs. crystallographic protein interfaces: an overview of computational approaches for their classification. *Crystals*, v. 10, n. 2, p. 114, 2020.

135 KAVANAGH, K. L. et al. Medium- and short-chain dehydrogenase/reductase gene and protein families. *Cellular and Molecular Life Sciences*, v. 65, n. 24, p. 3895, 2008a.

136 PERSSON, B.; KALLBERG, Y. Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs). *Chemico Biological Interactions, Enzymology and Molecular Biology of Carbonyl Metabolism,* v. 202, n. 1, p. 111–115, 2013.

137 VAN OVERTVELDT, S. *et al.* A structural classification of carbohydrate epimerases: From mechanistic insights to practical applications. *Biotechnology Advances*, v. 33, n. 8, p. 1814–1828, 2015.

138 FILLING, C. *et al.* Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *Journal of Biological Chemistry*, v. 277, n. 28, p. 25677–25684, 2002.

139 PERSSON, B. *et al.* The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative. *Chemico Biological Interactions*, v. 178, n. 1–3, p. 94–98, 2009.

140 DA COSTA, M. *et al.* Structure-function relationships in NDP-sugar active SDR enzymes: Fingerprints for functional annotation and enzyme engineering. *Biotechnology Advances*, v. 48, p. 107705, 2021.

141 KAVANAGH, K. L. *et al.* Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cellular and Molecular Life Sciences*, v. 65, n. 24, p. 3895–3906, 2008b.

142 SOMERS, W. S.; STAHL, M. L.; SULLIVAN, F. X. GDP-fucose synthetase from Escherichia coli: structure of a unique member of the short-chain dehydrogenase/reductase family that catalyzes two distinct reactions at the same active site. *Structure*, v. 6, n. 12, p. 1601–1612, 1998.

143 WILLIAMS, G. J. *et al.* Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-Amino-4-deoxy-L-arabinose biosynthesis. *Journal of Biological Chemistry*, v. 280, n. 24, p. 23000–23008, 2005.

144 THODEN, J. B.; FREY, P. A.; HOLDEN, H. M. Molecular structure of the NADH/UDPglucose abortive complex of udp-galactose 4-epimerase from *Escherichia coli*: implications for the catalytic mechanism. *Biochemistry*, v. 35, n. 16, p. 5137–5144, 1996.

145 BEIS, K. *et al.* The structure of NADH in the enzyme dTDP-d-glucose dehydratase (RmlB). *Journal of the American Chemical Society*, v. 125, n. 39, p. 11872–11878, 2003.

146 WIERENGA, R. K.; DE MAEYER, M. C. H.; HOL, W. G. J. Interaction of pyrophosphate moieties with. Alpha-helixes in dinucleotide-binding proteins. *Biochemistry*, v. 24, n. 6, p. 1346–1357, 1985.

147 ROSANO, C. et al. Probing the catalytic mechanism of GDP-4-keto-6-deoxy-d-mannose epimerase/reductase by kinetic and crystallographic characterization of site-specific mutants11Edited by K. Nagai. *Journal of Molecular Biology*, v. 303, n. 1, p. 77–91, 2000.

148 WALLACE, A. C.; LASKOWSKI, R. A.; THORNTON, J. M. Ligplot: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering*, v. 8, n. 2, p. 127–134, 1995.

149 JEZ, J. M. *et al.* Comparative anatomy of the Aldo-keto reductase superfamily. *Biochemical Journal*, v. 326, Pt 3, p. 625–636, 1997.

150 PENNING, T. M. The Aldo-keto reductases (AKRs): overview. *Chemico Biological Interactions*, v. 234, p. 236–246, 2015. https://doi: 10.1016/j.cbi.2014.09.024.

151 KAVANAGH, K. L. *et al.* Structure of xylose reductase bound to NAD<sup>+</sup> and the basis for single and dual co-substrate specificity in family 2 aldo-keto reductases. *Biochemical Journal*, v. 373, n. Pt 2, p. 319–326, 2003.

152 KOZMA, E. *et al.* The crystal structure of rat liver AKR7A1. a dimeric member of the Aldo-keto reductase superfamily. *Journal of Biological Chemistry*, v. 277, n. 18, p. 16285–16293, 2002.

153 OBMOLOVA, G. *et al.* Crystal structure of the *Escherichia coli* Tas protein, an NADP(H)-dependent Aldo-keto reductase. *Proteins*, v. 53, n. 2, p. 323–325, 2003.

154 BENNETT, M. J. *et al.* Steroid recognition and regulation of hormone action: crystal structure of testosterone and NADP<sup>+</sup> bound to 3 alpha-hydroxysteroid/dihydrodiol dehydrogenase. *Structure*, v. 5, n. 6, p. 799–812, 1997.

155 BARSKI, O. A.; TIPPARAJU, S. M.; BHATNAGAR, A. The Aldo-keto reductase superfamily and its role in drug metabolism and detoxification. *Drug Metabolism Reviews*, v. 40, n. 4, p. 553–624, 2008.

156 SUJATHA, M. S.; SASIDHAR, Y. U.; BALAJI, P. V. Energetics of galactose- and glucose-aromatic amino acid interactions: implications for binding in galactose-specific proteins. *Protein Science:* a publication of the protein society, v. 13, n. 9, p. 2502–2514, 2004.

157 SCHIMMEYER, J.; BOCK, R.; MEYER, E. H. L-galactono-1,4-lactone dehydrogenase is an assembly factor of the membrane arm of mitochondrial complex I in Arabidopsis. *Plant Molecular Biology*, v. 90, p. 117–126, 2016. https://doi: 10.1007/s11103-015-0400-4

158 HUH, W.-K. et al. Characterisation of d-arabinono-1,4-lactone oxidase from Candida albicans ATCC 10231. European Journal of Biochemistry, v. 225, n. 3, p. 1073–1079, 1994.

159 SALUSJÄRVI, T.; KALKKINEN, N.; MIASNIKOV, A. N. Cloning and characterization of gluconolactone oxidase of penicillium cyaneo-fulvum atcc 10431 and evaluation of its use for production of d-erythorbic acid in recombinant *Pichia pastoris*. *Applied and Environmental Microbiology*, v. 70, n. 9, p. 5503–5510, 2004.

160 ELICEIRI, G. L.; LAI, E. K.; MCCAY, P. B. Gulonolactone oxidase. solubilization, properties, and partial purification. *Journal of Biological Chemistry*, v. 244, n. 10, p. 2641–2645, 1969.

161 SUGISAWA, T. *et al.* Isolation and characterization of a new vitamin c producing enzyme (l-gulono-γ-lactone dehydrogenase) of bacterial origin. *Bioscience, Biotechnology, and Biochemistry*, v. 59, n. 2, p. 190–196, 1995.

162 BA, W.; D, C. *Mycobacterium tuberculosis* possesses a functional enzyme for the synthesis of vitamin C, L-gulono-1,4-lactone dehydrogenase. *FEBS Journal*, v. 273, n. 19, 2006. https://doi: 10.1111/j.1742-4658.2006.05443.x

163 LEFERINK, N. G. H.; VAN DEN BERG, W. A. M.; VAN BERKEL, W. J. H. l-Galactono-gamma-lactone dehydrogenase from *Arabidopsis thaliana*, a flavoprotein involved in vitamin C biosynthesis. *FEBS Journal*, v. 275, n. 4, p. 713–726, 2008b.

164 ÔBA, K. *et al.* L-galactono- $\gamma$ -lactone dehydrogenase: partial characterization, induction of activity and role in the synthesis of ascorbic acid in wounded white potato tuber tissue. *Plant and Cell Physiology*, v. 35, n. 3, p. 473–478, 1994.

165 YABUTA, Y. *et al.* Molecular characterization of tobacco mitochondrial l-galactono-γlactone dehydrogenase and its expression in *Escherichia coli*. *Plant and Cell Physiology*, v. 41, n. 6, p. 666–675, 2000.

166 CASTRO, J. C. *et al.* L-galactosa deshidrogenasa y l-gulono-1,4-lactona deshidrogenasa influyen en la biosíntesis de vitamina c en *Myrciaria dubia* (Kunth) McVaugh "camu-camu". *ECIPeru: Revista del Encuentro Científico Internacional,* v. 9, n. 2, p. 77–83, 2013.

167 EWING, T. A. *et al.* The VAO/PCMH flavoprotein family. *Archives of Biochemistry and Biophysics*, Flavoproteins: beyond the classical paradigms, v. 632, p. 104–117, 2017. https://doi: 10.1016/j.abb.2017.06.022

168 HOLM, L.; SANDER, C. Mapping the protein universe. *Science*, v. 273, n. 5275, p. 595–602, 1996.

169 FORNERIS, F. *et al.* Structural analysis of the catalytic mechanism and stereoselectivity in *Streptomyces coelicolor* alditol oxidase. *Biochemistry*, v. 47, n. 3, p. 978–985, 2008.

170 FRAAIJE, M. W. *et al.* Covalent flavinylation is essential for efficient redox catalysis in vanillyl-alcohol oxidase. *Journal of Biological Chemistry*, v. 274, n. 50, p. 35514–35520, 1999.

171 MEWIES, M.; MCINTIRE, W. S.; SCRUTTON, N. S. Covalent attachment of flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) to enzymes: the current state of affairs. *Protein Science:* a publication of the protein society, v. 7, n. 1, p. 7–20, 1998.